

FDZ
Forschungsdatenzentrum

der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt-
und Berufsforschung

07/2011
DE

FDZ-Datenreport

Dokumentation zu Arbeitsmarktdaten

Stichprobe der Integrierten Arbeitsmarktbiografien

Regionalfile 1975-2008 (SIAB-R 7508)

Matthias Dorner,
Marion König,
Stefan Seth



Bundesagentur für Arbeit

Stichprobe der Integrierten Arbeitsmarktbiografien Regionalfile 1975-2008 (SIAB-R 7508)

Matthias Dorner, Marion König, Stefan Seth (Institut für Arbeitsmarkt- und
Berufsforschung)

Datensatz: Version 1 (SIAB-R 7508 v1)

Dokumentation: Version 1 (SIAB-R_7508_DE_v1_dok1)

Die FDZ-Datenreporte beschreiben die Daten des FDZ im Detail. Diese Reihe hat somit eine doppelte Funktion: zum einen stellen Nutzerinnen und Nutzer fest, ob die angebotenen Daten für das Forschungsvorhaben geeignet sind, zum anderen dienen sie zur Vorbereitung der Auswertungen.

FDZ-Datenreporte (FDZ data reports) describe FDZ data in detail. As a result, this series of reports has a dual function: on the one hand, those using the reports can ascertain whether the data offered is suitable for their research task; on the other hand, the data can be used to prepare evaluations.

Zusammenfassung

Dieser FDZ-Datenreport beschreibt das „Regionalfile“ der Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2008 (SIAB-R 7508). Der Datensatz stellt die faktisch anonymisierte Version der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) dar und wird nach Abschluss eines Nutzungsvertrages mit dem IAB an Forschungseinrichtungen übermittelt.

Abstract

This FDZ data report describes the “Regionalfile” of the Sample of Integrated Labour Market Biographies 1975-2008 (SIAB-R 7508). This dataset represents the factual anonymous version of the Sample of Integrated Labour Market Biographies (SIAB) and may be delivered to scientific research institutions after concluding a use agreement with the IAB.

Keywords: German administrative micro data, labour market data, data manual, factual anonymized data, Scientific Use File

Disclaimer

Für ihre Mitarbeit und Unterstützung danken wir den Kolleginnen und Kollegen des Forschungsdatenzentrums (FDZ), die sich in der Arbeitsgruppe zur Stichprobe der Integrierten Arbeitsmarktbiografien beteiligt haben und wertvolle Anregungen und Kommentare für die Erstellung des Datensatzes gegeben haben.

Für den vorliegenden Datenreport wurden Teile des Datenreports der schwach anonymisierten Version der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) (FDZ Datenreport 01/2010) übernommen. Darüber hinaus sind einzelne Passagen an IAB-interne Datendokumentation der Rohdaten von IAB-ITM angelehnt.

Die Verantwortung für den Inhalt und mögliche Fehler verbleibt einzig bei den Autoren des vorliegenden Datenreports.

Inhaltsverzeichnis

1.	Einleitung und Kurzbeschreibung	6
1.1	Einleitung.....	6
1.2	Datennutzung.....	7
1.3	Veränderungen zu früheren Versionen.....	7
1.4	Kurzbeschreibung	9
1.5	Merkmalsliste	10
1.6	Mengengerüst.....	12
2	Datenquellen.....	12
2.1	Beschäftigten-Historik (BeH)	13
2.2	Leistungsempfänger-Historik (LeH).....	14
2.3	Leistungs-Historiken Grundsicherung (LHG/XLHG)	14
2.4	Arbeitsuchenden-Historiken (ASU/XASU).....	16
3	Datenaufbereitung, Stichprobe und Anonymisierung.....	16
3.1	Bereinigungen der Ausgangsdaten für die SIAB-R 7508.....	16
3.1.1	Beschäftigten-Historik (BeH)	16
3.1.2	Leistungsempfänger-Historik (LeH)	17
3.1.3	Arbeitsuchenden-Historik (X)ASU.....	17
3.1.4	Leistungs-Historiken Grundsicherung (X)LHG	18
3.1.5	Bereinigungen der IEB-Gesamtdatensatz.....	18
3.2	Stichprobendesign.....	19
3.3	Anonymisierung der SIAB-R 7508.....	19
3.3.1	Zusammenfassung von Quellen	20
3.3.2	Gültigkeitszeiträume der Quellen.....	20
3.3.3	Zensierung von Altersangaben	23
3.3.4	Aggregation und Löschungen von Variablen	23
3.3.5	Auswirkungen der Anonymisierung	25
3.4	Definition fehlender, ungültiger sowie anonymisierter Werte.....	26
3.5	Episodensplitting	26
4	Datenqualität und Probleme	28
4.1	Beschäftigten-Historik (BeH)	28
4.2	Leistungsempfänger-Historik (LeH).....	29
4.3	Leistungs-Historiken Grundsicherung (LHG/XLHG)	29
4.4	Arbeitsuchenden-Historiken (ASU/XASU).....	30
5	Merkmalsbeschreibungen.....	31
5.1	Identifikatoren.....	31
5.1.1	Systemfreie Personennummer (persnr)	31
5.1.2	Betriebsnummernzähler (bnn)	31

5.2	<i>Gültigkeitszeitraum</i>	32
5.2.1	Beginndatum Originalsatz (begorig)	32
5.2.2	Enddatum Originalsatz (endorig)	32
5.2.3	Beginndatum der Episode (begepi)	33
5.2.4	Enddatum der Episode (endeppi)	33
5.3	<i>Informationen zur Person</i>	33
5.3.1	Geschlecht (frau)	33
5.3.2	Geburtsjahr (gebjahr)	34
5.3.3	Staatsangehörigkeit (deutsch).....	34
5.3.4	Ausbildung (bild)	34
5.3.5	Schulausbildung (schbild)	36
5.4	<i>Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche</i>	36
5.4.1	Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund/ Abmeldegrund (grund_gr) ...	36
5.4.2	Tagesentgelt, täglicher Leistungssatz (tentgelt_gr)	38
5.4.3	Beruf – ausgeübte Tätigkeit (beruf_gr)	39
5.4.4	Stellung im Beruf und Arbeitszeit (stib)	40
5.4.5	Erwerbsstatus (erwstat_gr)	41
5.4.6	Beschäftigungsdauer vor dem Alter von 17 (tage_jung)	42
5.4.7	Beschäftigungsdauer nach dem Alter von 62 (tage_alt).....	42
5.5	<i>Betriebsmerkmale</i>	42
5.5.1	Wirtschaftszweig 73 – Wirtschaftsgruppe 73 (w73_gen_gr).....	42
5.5.2	Wirtschaftszweig 03 – Wirtschaftsgruppe 03 (w03_gr).....	43
5.6	<i>Ortsangaben</i>	44
5.6.1	Arbeitsort Region (ao_region)	44
5.6.2	Pendlerstatus (pendler)	44
5.7	<i>Generierte technische Merkmale</i>	45
5.7.1	Quelle des Satzes (quelle_gr)	45
5.7.2	Satzzähler Konto (spell)	45
5.7.3	Satzzähler pro Episode (level2)	46
5.7.4	Satzzähler pro Episode und Quelle (level1)	46
6	Literaturverzeichnis	48
7	Anhang	50
7.1	<i>Unterschiede in der Anonymisierung zentraler Merkmale im Vergleich zur IABS-R04</i>	50
7.2	<i>Zusätzliche Bereinigungsschritte im IABS-Regionalfile (IABS-R04) im Vergleich zum SIAB-Regionalfile (SIAB-R 7508)</i>	50
7.3	<i>Übersicht Datumsanpassungen und Löschungen</i>	53
7.4	<i>Vergleich SIAB 7508 und SIAB-Regionalfile 7508</i>	54
7.5	<i>Umschlüsselungstabellen vergrößerter Merkmale im Vergleich zur schwach anonymisierten Version der SIAB 7508</i>	55
7.6	<i>Auszählungen</i>	72
8	Abkürzungsverzeichnis	73

1. Einleitung und Kurzbeschreibung

1.1 Einleitung

Das Regionalfile der Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2008 (im Folgenden SIAB-R 7508) ist ein faktisch anonymisiertes¹ Scientific Use File der Stichprobe der Integrierten Arbeitsmarktbiografien 1975-2008 (SIAB 7508)². Das Regionalfile als spezifischer Typ von Scientific Use Files zeichnet sich durch die Erhaltung einer vergleichsweise tiefen regionalen Gliederung der Daten aus, bietet jedoch im Vergleich zur schwach anonymisierten Ausgangsdatei aus datenschutzrechtlichen Gründen lediglich ein eingeschränkteres Merkmalsspektrum auf Personen- bzw. Betriebsebene. Insgesamt umfasst das SIAB-Regionalfile 1975-2008 die Erwerbsverläufe von 1.515.463 Individuen, deren Erwerbsbiografien in 34.862.777 Datenzeilen dokumentiert sind.

Die Ausgangsdatei für das SIAB-Regionalfile bildet die schwach anonymisierte Version der SIAB 7508, eine 2%-Stichprobe aus den Integrierten Erwerbsbiografien (IEB) des Instituts für Arbeitsmarkt- und Berufsforschung (IAB).

Die IEB bestehen aus der Gesamtheit der Personen, die im Beobachtungszeitraum zwischen 1975 und 2008 mindestens einmal einen der folgenden Zustände aufweisen:

- Sozialversicherungspflichtige Beschäftigung (erfasst ab 1975)
- Geringfügige Beschäftigung (erfasst ab 1999)
- Bezug von Leistungen nach dem Rechtskreis SGB III (erfasst ab 1975) oder SGB II (erfasst ab 2005)
- Bei der Bundesagentur für Arbeit (BA) als arbeitssuchend gemeldet (erfasst ab 2000)
- (Geplante) Teilnahme an arbeitsmarktpolitischer Maßnahme (erfasst ab 2000)

Diese aus unterschiedlichen administrativen Datenquellen stammenden Informationen werden in den IEB zusammengeführt und die Zustände jeweils tagesgenau abgebildet.

Während die Beschäftigten-Historik (BeH) Ursprung der Informationen zu sozialversicherungspflichtiger und geringfügiger Beschäftigung ist, wird der Bezug von Leistungen nach den Rechtskreisen des SGB III und des SGB II in der Leistungsempfänger-Historik (LeH) bzw. in den Leistungs-Historiken Grundsicherung (LHG und XLHG) erfasst. Die Arbeitssuchenden-Historiken (ASU und XASU) stellen die Datenquelle für die durch die BA erfassten Zeiten der Arbeitsuche dar, während die (geplante) Teilnahme an arbeitsmarktpolitischen Maßnahmen in der Maßnahme-Teilnahme-Historik (MTH) erfasst werden. Mit Ausnahme der (geplanten) Teilnahme an arbeitsmarktpolitischen Maßnahmen aus der MTH sind alle Da-

¹ Zum Konzept der faktischen Anonymität siehe Abschnitt 3.3.

² Die schwach anonymisierte Version SIAB 7508 wird in Dorner et al. (2010) beschrieben.

tenquellen der IEB sowohl im Regionalfile der SIAB als auch in der schwach anonymisierten Version der SIAB enthalten.

Die SIAB ersetzt die IAB-Beschäftigtenstichprobe (IABS) und die Stichprobe der Integrierten Erwerbsbiographien (IEBS). Durch die Zusammenführung der Datenquellen von IABS und IEBS kann nun mit der SIAB ein Datensatz angeboten werden, der die Vorteile der IABS (langer Beobachtungszeitraum) mit den Vorteilen der IEBS (vollständige Erwerbsbiografien) verbindet. Die SIAB wird im Forschungsdatenzentrum (FDZ) der Bundesagentur für Arbeit (BA) im IAB produziert.

Der vorliegende Datenreport beschreibt die Erstellung des SIAB-Regionalfiles sowie die im Scientific Use File enthaltenen Merkmale. Im Weiteren gliedert sich der Datenreport wie folgt: neben der Einleitung beinhaltet Kapitel 1 Informationen zum Datenzugang sowie eine Kurzbeschreibung der Daten, das Mengengerüst und eine Merkmalsliste. Eine Beschreibung der einzelnen Datenquellen findet sich in Kapitel 2. Die Datenaufbereitung sowie das Anonymisierungskonzept des Scientific Use Files werden in Kapitel 3 beschrieben. Nachdem in Kapitel 4 die Datenqualität diskutiert wird, sind die einzelnen im Datensatz enthaltenen Merkmale abschließend in Kapitel 5 dargestellt.

1.2 Datennutzung

Das SIAB-Regionalfile wird auf Antrag Forscherinnen und Forschern zeitlich begrenzt für die Durchführung eines inhaltlich festgelegten Projekts zur Verarbeitung in Räumen ihrer Forschungseinrichtung zur Verfügung gestellt (§ 282 Abs. 7 SGB III).

Für die Datennutzung ist zunächst ein Projektantrag beim FDZ einzureichen. Nach Genehmigung durch das FDZ wird ein projektspezifischer Nutzungsvertrag mit der Institution des Forschers geschlossen.

Details zu verschiedenen Datenzugangswegen, zu den Voraussetzungen der Datennutzung sowie zur Beantragung des Datensatzes können der FDZ-Homepage unter <http://fdz.iab.de/> entnommen werden.

1.3 Veränderungen zu früheren Versionen

Die Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) ersetzt die IAB Beschäftigtenstichproben (IABS) sowie die Stichprobe aus den Integrierten Erwerbsbiographien (IEBS).

Da die beiden Vorgängerdatensätze jeweils auch die Datenbasis für eigene Scientific Use Files und damit die Vorgänger des SIAB-Regionalfiles bilden, werden im Folgenden kurz die wichtigsten Unterschiede zwischen diesen Datensätzen dargestellt.

Scientific Use Files des FDZ (Auswahl)			
	SIAB	IABS	IEBS
Typ des Scientific Use Files	Regionalfile ³ (SIAB- R 7508)	Regionalfile (IABS- R 7504)	Basisfile ⁴
Datenbasis	Integrierte Erwerbsbiografien (IEB)	Beschäftigten- und Leistungsempfänger-Historik (BLH)	Integrierte Erwerbsbiografien (IEB)
Quellen	Beschäftigung, Leistungsempfang, Arbeitsuche	Beschäftigung, Leistungsempfang	Beschäftigung, Leistungsempfang, Arbeitsuche, Maßnahmen
Beobachtungszeitraum Beschäftigung	1975-2008	1975-2004	1993-2003
Erhebungsdesign	2% Zufallsstichprobe der Personen in den IEB	Geschichtete 2% Zufallsstichprobe der Beschäftigten der BLH	Geburtsstichprobe der Personen aus IEB, $8/365 \approx 2,19\%$
Jahresmeldungen mit Personengruppe 101 und Entgelt == 0	Nicht enthalten	enthalten	Nicht enthalten
Aufbereitung der Basisdaten	Marginale Korrekturen der IEB	Umfangreiche Korrekturen, Ergänzungsverfahren	Marginale Korrekturen der IEB

Tabelle 1: Unterschiede zwischen den Scientific Use Files der SIAB, IABS und IEBS

Anders als in der SIAB-R 7508 und dem IEBS-SUF wurden für das IABS-Regionalfile vor der Anonymisierung eine Reihe zusätzlicher Bereinigungsschritte durchgeführt. Dem Anhang des Datenreports kann eine Beschreibung dieser Korrekturen und Bereinigungen in den Vorgängerdaten der IABS entnommen werden, die für die SIAB nicht mehr durchgeführt wurden (vgl. Abschnitte [7.1](#), [7.2](#)).

³ Mit einem sog. „Regionalfile“ bezeichnen wir ein Scientific Use File, welches eine vergleichsweise tiefe regionale Gliederung der Daten ausweist, jedoch aus datenschutzrechtlichen Gründen ein nur stark reduziertes Informationsspektrum auf Personen- und Betriebsebene umfasst.

⁴ Mit einem sog. „Basisfile“ bezeichnen wir ein Scientific Use File, das detaillierte personen- bzw. betriebsbezogene Informationen ausweist, jedoch aus Anonymitätsgründen kein bzw. nur sehr grob gegliederte räumliche Einheiten beinhaltet.

1.4 Kurzbeschreibung

Aktuelle Datenversion	Stichprobe der Integrierten Arbeitsmarktbiografien - Regionalfile 1975 – 2008 (SIAB-R 7508), Version 1 Kurzform: SIAB-R 7508 v1
Aktuelle Version der Dokumentation	Stichprobe der Integrierten Arbeitsmarktbiografien - Regionalfile (SIAB-R) 1975 – 2008, Version 1 Kurzform: SIAB-R_7508_DE_v1_dok1
Art und Umfang der Daten	Faktisch anonymisierte Daten (Scientific Use File), 2%-Zufallsauswahl von Personenkonten aus den Integrierten Erwerbsbiografien (IEB) des IAB
Kurzbeschreibung des Inhalts	Beschäftigten-Historik (BeH): Jahres- und Abmeldungen zur Sozialversicherung für sozialversicherungspflichtig und geringfügig Beschäftigte Leistungsempfänger-Historik (LeH): Informationen zu Leistungsbezug nach SGB III für Bezieher von Arbeitslosengeld, Arbeitslosenhilfe und Unterhaltsgeld Leistungs-Historiken Grundsicherung aus A2LL(LHG) sowie aus XSozial-BA-SGB II (XLHG): Daten zu Personen, die Leistungen zur Sicherung des Lebensunterhalts nach SGB II beziehen (ARGE, getrennte Trägerschaft, sowie zugelassene kommunale Träger) Arbeitsuchenden-Historik (ASU) sowie aus XSozial-BA-SGB II (XASU): Informationen zur Arbeitsuche (Bewerberangebote) inklusive den Informationen zur Arbeitsuche, die über den Übermittlungsstandard XSozial-BA-SGB II von zugelassenen kommunalen Trägern oder getrennten Trägern an die BA gemeldet werden
Merkmale mit regionaler Relevanz	333 kreisfreie Städte/Landkreise bzw. Kreisregionen auf Basis des Arbeitsorts, kategoriales Merkmal zum Pendlerstatus/Wohnort
Gebietsstand Datum	31.12.2008
Gebietsstand Typ	BeH, LeH, (X)ASU, (X)LHG: gebietsstandsbereinigt
Zeitraum	Der Zeitraum hängt von der Datenquelle ab. BeH 1.1.1975 - 31.12.2008 (2007: 18-Monatsdatei; 2008: 12-Monatsdatei) LeH 1.1.1975 - 31.12.2008 (X)ASU 1.1.2000 - 31.12.2004 und 1.1.2007 - 31.12.2008 (X)LHG 1.1.2007 - 31.12.2008
Frequenz der Aktualisierung	Laufend
Untersuchungseinheit	Sozialversicherungspflichtig Beschäftigte (ab 1999 auch geringfügig Beschäftigte), Leistungsempfänger, Arbeitsuchende.
Fallzahlen	1.515.463 Personen, 34.862.777 Sätze.

Bereinigungen - Kurzbeschreibung	Die Daten des SIAB-Regionalfiler durchlaufen nach der Ziehung aus den IEB weitere Aufbereitungs- und Anonymisierungsschritte wie Umkodierungen einzelner Merkmale, Vergrößerungen bzw. auch Löschungen von Variablen und Zeiträumen.
Besonderheiten - Kurzbeschreibung	Die Daten im SIAB-Regionalfiler sind episodengesplittet. Dabei werden sich teilweise überschneidende Datensätze innerhalb eines Personenkontos gesplittet, sodass nur überschneidungsfreie oder vollständig parallele Zeiträume vorkommen.
Dateiformat, Dateigröße	Stata (circa 1,4 GB), SPSS (circa 2,8 GB)
Datenzugang	Übermittlung der Daten ausschließlich an öffentliche Forschungseinrichtungen bzw. Einrichtungen, die nachweislich unabhängige wissenschaftliche Forschung betreiben.
Anonymisierungsgrad	Faktisch anonyme Daten (Scientific Use File)
Zitierweise	Stichprobe der integrierten Arbeitsmarktbiografien - Regionalfiler 1975-2008 (SIAB-R 7508), Nürnberg 2011 Kurzform: SIAB-R 7508 v1

1.5 Merkmalsliste

Die Merkmalsübersicht listet die Variablennamen und die längeren Merkmalsbezeichnungen auf. Weiterhin bietet sie eine Übersicht darüber, welche Merkmale für welche Datenquellen gefüllt sind:

	Merkmal ist für die Quelle gefüllt. Füllgrad immer > 0,85
	Merkmal ist für die Quelle gefüllt. Niedriger bzw. wechselnder Füllgrad, siehe Merkmalsbeschreibung und Häufigkeitsauszählung
	Merkmal ist für diese Quelle nicht gefüllt. Füllgrad immer < 0,05

Beispiel: Der Schulabschluss aus dem Merkmal Schulbildung (schbild) ist lediglich für (X)ASU-Sätze gefüllt, die Sätze der anderen Datenquellen enthalten in dieser Variablen missing (.n). Eine weitere Besonderheit der SIAB liegt darin, dass einige Merkmale je nach Datenquelle unterschiedliche Inhalte haben. So enthält der Erwerbsstatus (erwstat_gr) für BeH-Sätze die Personengruppe des Meldeverfahrens, für LeH-Sätze die Leistungsart, für (X)LHG-Sätze den Status SGB-II und für (X)ASU Sätze den Arbeitsuche-Status. Nicht bei jedem Merkmal sind diese Unterschiede direkt aus dem Variablennamen ersichtlich.

Merkmalstyp	Merkmal	Seite	1 BeH	2 LeH	3 (X)ASU	4 (X)LHG
			1975- 2008	1975- 2008	2000-2004 2007-2008	2007- 2008
Identifikatoren	persnr	31				
	bnn	31				
Gültigkeitszeitraum/ Datumsangaben	begorig	32				
	endorig	32				
	begepi	33				
	endeipi	33				
Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche	quelle_gr	45				
	erwstat_gr	41				
	grund_gr	36				
	tentgelt_gr	38				
Informationen zur Person	deutsch	34				
	gebjahr	34				
	frau	33				
	bild	34				
	beruf_gr	39				
	stib	40				
	schbild	36				
	tage_jung	42				
	tage_alt	42				
(Generierte) Betriebs- merkmale	w73_gen_gr	42				
	w03_gr	43				
Ortsangaben	ao_region	44				
	pendler	44				
Generierte technische Merkmale	spell	45				
	level1	46				
	level2	46				

Tabelle 2: Merkmalsliste und merkmalspezifische Füllgrade nach Quellen

1.6 Mengengerüst

Fallzahlen der SIAB-R 7508	vor Splitting	nach Splitting
Beschäftigung (BeH)	23.642.269	25.825.849
Leistungsempfang SGB III (LeH)	3.530.492	4.671.200
Arbeitsuche, auch SGB II (ASU/XASU)	2.185.067	3.588.475
Leistungsempfang SGB II (LHG/XLHG)	324.396	777.253
Sätze insgesamt	29.682.224	34.862.777
Personen	1.515.463	

Tabelle 3: Mengengerüst SIAB-Regionalfiler 7508

2 Datenquellen

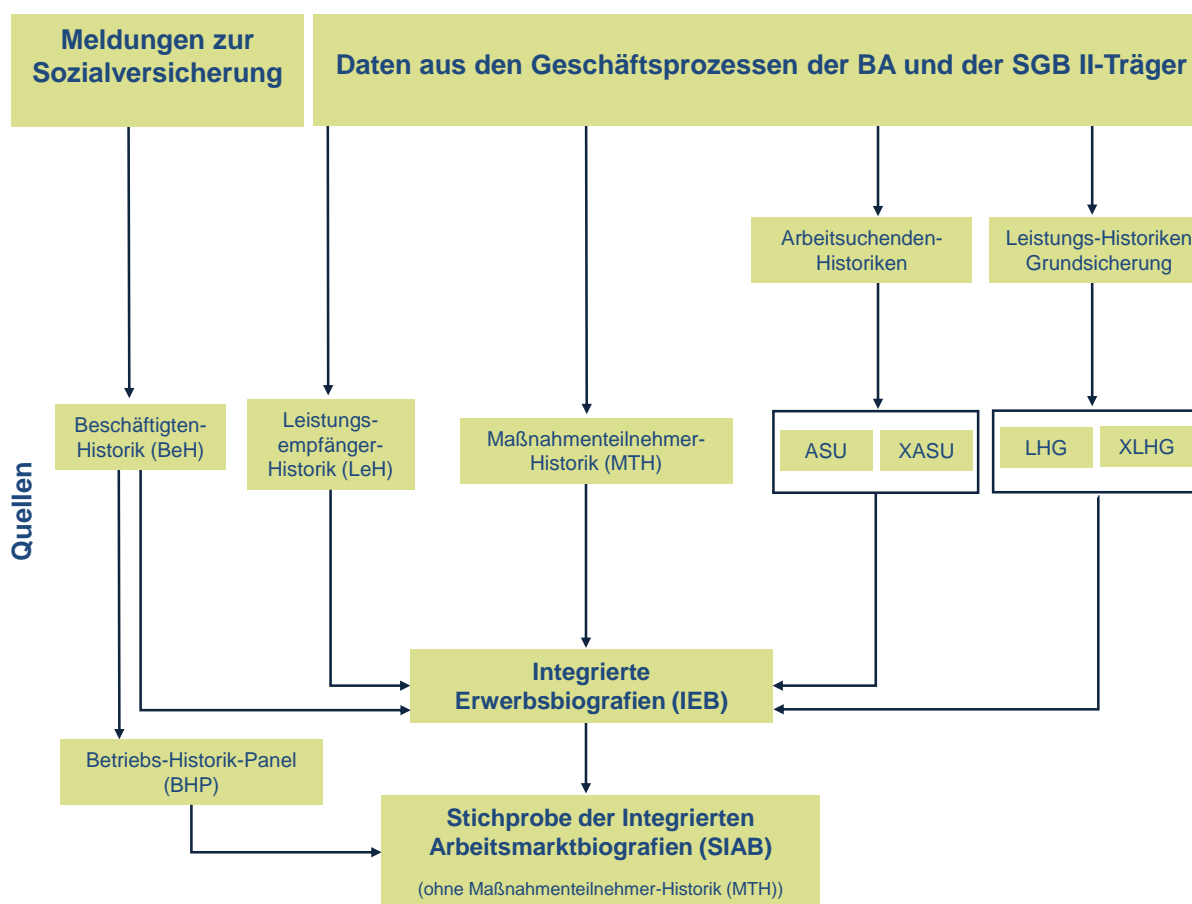


Abbildung 1: Datenquellen der Integrierten Erwerbsbiografien (IEB)

Die SIAB wurde als 2%-Stichprobe aus den Integrierten Erwerbsbiografien (IEB) des IAB gezogen. Diese vereint in sich wiederum Daten aus fünf verschiedenen Datenquellen, die jeweils Informationen aus unterschiedlichen Verfahren enthalten können.

In die SIAB werden Datensätze aus allen dargestellten Quellen mit Ausnahme der (geplanten) Teilnahmen an arbeitsmarktpolitischen Maßnahmen (MTH) aus den IEB übernommen. Diese Datenquellen werden im Folgenden näher beschrieben.

2.1 Beschäftigten-Historik (BeH)

Die Quelle für Informationen über Beschäftigung ist die Beschäftigten-Historik des IAB. Die rechtliche Grundlage dieser Daten ist das mit Wirkung vom 1. Januar 1973 eingeführte (und seit dem 1. Januar 1991 auf Ostdeutschland ausgedehnte) integrierte Meldeverfahren zur Kranken-, Renten- und Arbeitslosenversicherung, das durch das Kürzel DEÜV (früher DEVO/DÜVO) bezeichnet ist (vgl. für weitere Details: Bender et al. 1996, S. 4 ff.; Wermter/Cramer 1988). Es verlangt von den Arbeitgebern Meldungen für alle sozialversicherungspflichtig beschäftigten Arbeitnehmer an die zuständigen Sozialversicherungsträger. Die BeH erfasst alle Arbeiter und Angestellten sowie alle Auszubildenden, soweit sie nicht von der Sozialversicherungspflicht befreit sind. Dies bedeutet, dass Beamte, Selbstständige und ordentliche Studierende (vgl. Cramer 1985) grundsätzlich nicht in der BeH erfasst werden. Seit der Änderung des Meldeverfahrens zum 1. Januar 1999 werden auch geringfügig Beschäftigte und mithelfende Familienangehörige erfasst (enthalten erst ab 1. April 1999).

Jedes Jahr, in dem ein Beschäftigungsverhältnis für eine Person existiert, ist durch mindestens eine Meldung abgebildet. Die Daten werden von den Krankenkassen aufgenommen, von der Bundesagentur für Arbeit in einer laufenden Datei gesammelt und anschließend vom IAB in seine Historik-Datei integriert.

Die Merkmale des Regionalfiles der SIAB-Daten, die sich auf Betriebe beziehen, werden dem Betriebs-Historik-Panel (BHP) entnommen, das ebenfalls auf der BeH basiert (vgl. Hethey-Maier, Seth 2010). Im Scientific Use File enthalten sind dabei folgende auf dem BHP basierende und angepasste Merkmale:

- Arbeitsregion (ao_region)
- Wirtschaftszweig 73 (extrapoliert bis 2008) (w73_gen_gr)
- Wirtschaftszweig 03 (w03_gr)

Zwei Dinge müssen bei den Betriebsinformationen berücksichtigt werden:

1. Die Merkmale zu den Betrieben werden zum 30.6. eines Jahres erfasst und auf Basis der Personendaten für die entsprechende Betriebsnummer aggregiert. Die Betriebs-

merkmale sind deshalb, im Unterschied zu den Personenangaben, keine zeitraumbezogenen Daten, sondern nur am 30.6. jedes Jahres exakt gültig⁵.

2. Zum anderen wird das BHP direkt aus der BeH erstellt, die SIAB Personen-Datei jedoch aus den IEB (siehe Abbildung 1), in die nicht alle Datensätze der BeH übernommen werden.

2.2 Leistungsempfänger-Historik (LeH)

Die Leistungsempfänger-Historik des IAB erfasst Zeiträume, in denen Personen Lohnersatzleistungen von der Bundesagentur für Arbeit (BA) beziehen. Die Leistungen beinhalten Arbeitslosengeld, Arbeitslosenhilfe und Unterhaltsgeld, also ausschließlich Leistungen nach dem Rechtskreis SGB III. Da der Leistungsanspruch von der Erfüllung gesetzlicher Voraussetzungen abhängt, werden Arbeitslosigkeitsperioden, in denen die Voraussetzungen nicht vorliegen (z.B. fehlende Bedürftigkeit im Falle der Arbeitslosenhilfe oder Nichterfüllung der Anwartschaftszeit beim Arbeitslosengeld) nicht in der Leistungsempfänger-Historik berichtet.

2.3 Leistungs-Historiken Grundsicherung (LHG/XLHG)

Während die LeH die Leistungen des Rechtskreises SGB III enthält, beinhalten LHG und XLHG Leistungsbezug nach SGB II. Dies umfasst sowohl Leistungen zur Sicherung des Lebensunterhalts (z. B. ALG II) wie auch Zuschläge zum Arbeitslosengeld oder auch zusätzliche Leistungen. Im Gegensatz zu den Leistungen aus dem Rechtskreis des SGB III ist die Bundesagentur für Arbeit nicht mehr zwingend alleiniger Träger der Leistungen. Es wird demnach grundsätzlich unterschieden nach den drei möglichen Trägerarten:

- Arbeitsgemeinschaften (ARGE) aus gemeinsamer Aufgabenwahrnehmung von BA und Kommune.
- Bei getrennter Trägerschaft bleibt die Aufgabenteilung zwischen BA und Kommune bestehen⁶.

⁵ Extrembeispiel: Es liegt eine Beschäftigungsmeldung vom 1.1. bis zum 30.5.2006 vor, der Betrieb geht im Juni 2006 Konkurs. Dann gibt es im BHP zu diesem Betrieb für 2006 keine Informationen.

⁶ Die Kommune übernimmt die Kosten der Unterkunft und Heizung (§ 22 SGB II) sowie die abweichend zu erbringenden Leistungen (§ 23 (3) SGB II) und die flankierenden Leistungen nach §16 (2) Satz 2 Nr. 1 - 4 SGB II. Von der BA werden dagegen Regelleistungen, Sozialversicherungsbeiträge und Eingliederungsleistungen (SGB III und SGB II) sowie spezifische Leistungen ohne die oben genannten flankierenden Leistungen erbracht.

- Bei zugelassenen kommunalen Trägern, die auch als optierende Kommunen oder Optionskommunen bezeichnet werden, übernimmt die Kommune Aufgaben aus dem Bereich des SGB II.

Die Leistungs-Historiken Grundsicherung umfassen zwei Quellen, die sich im Übermittlungsverfahren der Daten unterscheiden: Die Daten der „Leistungs-Historik Grundsicherung aus A2LL“ (LHG) stammen aus dem BA-Verfahren A2LL, während die „Leistungs-Historik-Grundsicherung aus XLHG-BA-SGBII“ (XLHG) Daten erfasst, welche nach dem Übermittlungsstandard X-Sozial gemeldet wurden.

In der Regel findet A2LL bei ARGE n Verwendung, während der Standard X-Sozial von den optierenden Kommunen genutzt wird. Bei Kommunen mit getrennter Trägerschaft kommen beide Verfahren zum Einsatz.

Die Ermittlung der Höhe des Leistungsbezugs erfolgt nicht auf individueller Ebene, sondern auf der Ebene von Bedarfsgemeinschaften. Eine Verknüpfung von Individuen mit SGB-II-Leistungsbezug auf Ebene der Bedarfsgemeinschaften ist in diesem Datensatz jedoch nicht möglich. In LHG und XLHG sind grundsätzlich alle gemeldeten Berechtigten nach § 7 SGB II, also erwerbsfähige Hilfebedürftige und ggf. ihre Bedarfsgemeinschaftsmitglieder enthalten.

Es sind jeweils überschneidungsfreie Zeiträume der Leistungsberechtigung einer Person in einer bestimmten Bedarfsgemeinschaft abgebildet. Die Sätze werden jedoch

- an transferrelevanten Geburtstagen der Mitglieder (14, 15, 18, 25 und 65) einer Bedarfsgemeinschaft,
- bei Veränderungen in der Zusammensetzung einer Bedarfsgemeinschaft (z.B. durch Zu-/Abgänge),
- bei Merkmalsänderungen des Bedarfsgemeinschafts-Kunden sowie
- bei Beginn und Ende eines Sanktionssachverhalts für Sätze ab dem 1. April 2006 gesplittet.

Im Hinblick auf Sanktionen ist zu beachten, dass Beginn und Ende eines Sanktionssachverhalts zwar zu einem Splitting der Sätze führen, jedoch anhand der Daten weder die Dauer, die Art noch der Zeitpunkt der Verhängung bzw. des Beginns der Sanktion festgestellt werden kann. Grund hierfür ist das Fehlen eines entsprechenden Merkmals oder einer Ausprägung, welches den Beginn, die Art oder die Dauer der Sanktion anzeigt.

2.4 Arbeitssuchenden-Historiken (ASU/XASU)

Daten über Arbeitssuchende werden in der Arbeitssuchenden-Historik (ASU) gespeichert. Die Quelle ASU enthält dabei Informationen zu Arbeitssuchenden, welche bei Arbeitsagenturen verzeichnet sind und schließt ab 2005 auch getrennte Trägerschaften mit ein.

In der Quelle XASU dagegen sind ab 2005 die Daten von arbeitssuchenden Personen im ALG-II-Bezug von zugelassenen kommunalen Trägern enthalten, welche nach dem X-Sozial-Standard gemeldet werden.

3 Datenaufbereitung, Stichprobe und Anonymisierung

3.1 Bereinigungen der Ausgangsdaten für die SIAB-R 7508

Die Integrierten Erwerbsbiografien des IAB (IEB) mit ihren verschiedenen Quelldateien bilden eine wichtige Datenbasis für eine Vielzahl vom FDZ der BA im IAB bereitgestellte Datensätze. Bis die IEB jedoch für die Erstellung der FDZ-Datenprodukte bereitsteht, durchläuft sie je nach Datensatz quellenspezifische Bereinigungen unterschiedlichen Umfangs.

Im Folgenden sollen diejenigen Aufbereitungen in den IEB dargestellt werden, die für die Erstellung der SIAB 7508 sowie des darauf beruhenden Scientific Use Files SIAB-R 7508 relevant sind.

3.1.1 Beschäftigten-Historik (BeH)

- Sätze mit einem Entgelt in Höhe von Null bzw. mit fehlendem Entgelt sowie der Ausprägung 101 beim Merkmal Personengruppe und dem Wert 50 beim Abgabegrund (Jahresmeldung) werden nicht in die IEB übernommen.
- Um einen im Zeitablauf möglichst konstanten Personenkreis zu erfassen, werden einige Personengruppen, deren Daten nicht für den gesamten Beobachtungszeitraum zur Verfügung stehen, ausgeschlossen. Konkret sind ab 1999 nur die Personengruppen 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 149, 201, 203 und 205 enthalten sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind ab 1999 unter anderem kurzfristig Beschäftigte, d.h. Personengruppen 110, 202 und 210.
- Geschlecht und Geburtsdatum werden aus der Versicherungsnummer übernommen. Zu beachten ist aber, dass diese Angaben quellenübergreifend harmonisiert werden.
- Der Gebietsstand für Arbeitsort und Wohnort wird zum Stand vom 31.12.2008 aktualisiert.

3.1.2 Leistungsempfänger-Historik (LeH)

- Aus der LeH werden für die IEB nur Sätze für den Zeitraum ab 1.1.1975 (Enddatum) bis zum aktuellen Rand übernommen.
- Zu beachten ist, dass bei noch laufendem Bezug das Enddatum dem Ende des Bewilligungszeitraums entspricht und nicht dem Ende des Bezugszeitraums. Im SIAB-Regionalfiler werden keine Angaben zum Ende des Bewilligungszeitraums ausgewiesen, da zum 31.12.2008 noch gültige Leistungsempfangsmeldungen rechtszensiert werden (vgl. Abschnitt 3.3.2). Sätze mit leerem oder ungültigem Enddatum werden ausgeschlossen, da hier nicht davon ausgegangen werden kann, dass überhaupt eine Zahlung geleistet worden ist.
- Sätze ohne gültiges Beginndatum werden ausgeschlossen.
- Sätze mit einem Enddatum, das vor dem Beginndatum liegt, werden ausgeschlossen.
- Liegt das Enddatum für die Arbeitslosenhilfe einen Tag vor dem Beginndatum, so wird das Enddatum um ein Jahr verschoben.
- Sich überschneidende LeH-Sätze werden für die IEB entsprechend der folgenden Regeln bereinigt:
 - Alle Überschneidungen, bei denen in beiden Sätzen dieselbe Leistungsart und dieselbe Leistungshöhe steht, werden wie folgt bereinigt:
 - Wenn ein Satz vollständig innerhalb des Zeitraums eines anderen Satzes liegt, dann wird der Satz mit der kürzeren Dauer gelöscht. Bei gleicher Länge wird ein (zufällig ausgewählter) Satz gelöscht.
 - Bei Teilüberschneidungen wird das Enddatum des früher beginnenden Satzes korrigiert.
 - Überschneidungen zwischen Sätzen mit **unterschiedlicher Leistungsart** und/oder unterschiedlicher Leistungshöhe werden nur dann bereinigt, wenn es sich um eine Überschneidung von genau einem Tag handelt und die Originaldauer beider Sätze größer als ein Tag ist. In diesen Fällen wird das Enddatum des früher beginnenden Satzes um einen Tag zurückgesetzt.

3.1.3 Arbeitsuchenden-Historik (X)ASU

- Die aus der ASU verwendeten Sätze umfassen den Zeitraum von Januar 2000 bis Dezember 2008. Die registrierten Zeiten der Arbeitsuche der einzelnen Personen werden als vollständig angesehen.
- In der Datenbasis ASU gibt es nur die Unterscheidung zwischen Sätzen mit Status "arbeitslos" und "arbeitsuchend". In der IEB wird jedoch zwischen "arbeitslos", "krank" und

"arbeitsuchend" unterschieden. Der Erwerbsstatus "krank" wird IEB-Sätzen zugeordnet, wenn in der ASU-Datenbasis

- ein Vorgängersatz mit Status „arbeitslos“ existiert, lückenlos anschließt und den Abgangsgrund "krank" hat und
 - ein Nachfolgesatz mit Status „arbeitslos“ existiert und ebenfalls lückenlos anschließt
 - und der Satz selbst nicht den Status "arbeitslos" sondern den Status "arbeitsuchend" hat.
- In der XASU werden im Gegensatz zur ASU Krankheitsperioden nicht berücksichtigt, da in den XASU-Daten keine Informationen zu Krankheiten vorliegen.
- Sätze mit einem Enddatum, das kleiner als das Beginndatum ist, werden ausgefiltert.

3.1.4 Leistungs-Historiken Grundsicherung (X)LHG

- Es werden nur Sätze erwerbsfähiger Personen und von Personen unter 64 Jahren eingelesen.
- Im Sozialgesetzbuch festgelegte Änderungen in der Struktur von Bedarfsgemeinschaften haben ein Splitting von einzelnen Sätzen zur Folge. Auf diese Weise lässt sich unter Umständen auf datenschutzrechtlich als sensibel eingestufte personenbezogene Informationen wie das exakte Geburtsdatum eines Transferempfängers schließen. Aus diesem Grund werden die betroffenen Sätze bzw. die zugehörigen Datumsangaben nach einem festgelegten Anonymisierungsschema umgesetzt.

3.1.5 Bereinigungen der IEB-Gesamtdatetei

Nachdem die Rohdaten bereinigt und zur Grundgesamtheit der IEB zusammengefasst wurden, erfolgt eine quellenübergreifende Bereinigung der Gesamtdatetei der IEB. Dabei werden folgende Arbeitsschritte durchgeführt:

- Datensätze, bei denen das Alter kleiner 13 oder größer 75 ist, werden gelöscht.
- Datensätze, bei denen das Enddatum kleiner als das Beginndatum ist, werden gelöscht.
- Personenkonten mit einem Wechsel von Geschlecht oder Geburtsdatum werden nach festgelegten Prioritätsregeln bereinigt.

Die auf dieser Datenbasis erstellten Daten der schwach anonymisierten SIAB 7508 werden nicht mehr weiter bereinigt und entsprechen somit (nahezu) den Originaldaten der IEB. Es werden lediglich vereinzelte Umkodierungen einzelner Variablen oder aus Anonymisierungsgründen auch Vergrößerungen der Daten durchgeführt. Diese Daten bilden die Datengrund-

lage für das Scientific Use File der Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB-R 7508).

3.2 Stichprobendesign

Ausgangsbasis für das Scientific Use File ist die schwach anonymisierte Version der SIAB, die eine 2%-Stichprobe derjenigen Arbeitsmarktteilnehmer darstellt, für die ein Eintrag in den IEB vorliegt (vgl. Dorner et al. 2010). Entsprechend lassen sich in der Stichprobe mit dem Hochrechnungsfaktor 50 auch Fallzahlen der Grundgesamtheit schätzen. Dies gilt vor dem Hintergrund des Stichprobendesigns grundsätzlich für die im Datensatz enthaltenen Personenkreise aller Datenquellen. Je nach Fragestellung und der Auswahl der Merkmale sollte jedoch hinsichtlich der Repräsentativität zusätzlich der unterschiedliche Füllgrad von Merkmalen (vgl. Tabelle 2) innerhalb der einzelnen Quellen sowie deren spezifische Gültigkeitszeiträume beachtet werden.

Die Ziehung der SIAB-Beobachtungen aus den IEB ist vergleichbar mit der Ziehung der IEBS.⁷ Im Vergleich zur IAB-Beschäftigtenstichprobe (IABS) gibt es hingegen zwei Unterschiede hinsichtlich der Stichprobenziehung. Zum einen ist die IABS eine Beschäftigtenstichprobe, d.h. die Personen wurden aus den Beschäftigtenkonten der BeH gezogen, Leistungssätze aus der LeH wurden diesen Konten lediglich zugespielt. Es gibt in der IABS also keine Konten, die ausschließlich LeH-Sätze enthalten, während es in der SIAB sehr wohl Personen gibt, die in ihrem Konto keine einzige Beschäftigungsmeldung aufweisen. Zum anderen wird die SIAB im Gegensatz zur IABS nicht nach Ost/West und deutsch/nichtdeutsch geschichtet gezogen. Aufgrund der hohen Fallzahl sind diese vier Personenkreise jedoch auch ohne explizite Schichtung mit annähernd gleicher Auswahlwahrscheinlichkeit in der Stichprobe enthalten.

3.3 Anonymisierung der SIAB-R 7508

Die Stichprobe der integrierten Arbeitsmarktbiografien, die als Datengrundlage des SIAB-Regionalfiles dient, enthält eine Vielzahl personenbezogener Informationen zum Status einer Person auf dem Arbeitsmarkt sowie zusätzliche Betriebsinformationen bei Beschäftigungsmeldungen aus dem Meldeverfahren der Sozialversicherung. Diese im FDZ auf Mikroebene vorliegenden Sozialdaten (§ 67 SGB X) dürfen, so schreibt der Gesetzgeber vor, nur zu wissenschaftlichen Zwecken weitergegeben werden (§ 282 Abs. 7 SGB III, § 75 SGB X), wenn

⁷ Die Ziehung der IEBS aus den IEB erfolgt über eine zufällige Auswahl von acht Geburtstagen (vgl. Oberschachtsiek et al. 2008).

sowohl die personenbezogenen als auch die betriebsbezogenen Daten nach einer Prüfung als faktisch anonym angesehen werden können. Dies bedeutet, dass unverhältnismäßig viel Zeit, Kosten und Arbeit aufgewendet werden müssen, um einen Merkmalsträger im Datensatz zu deanonymisieren. Die Kosten einer Deanonymisierung müssen deren Nutzen also übersteigen.

Die Vorgaben von Müller et al. (1991) zur faktischen Anonymität von Mikrodaten bilden den maßgeblichen Orientierungsrahmen bei der Erstellung von Scientific Use Files am FDZ der BA im IAB (vgl. u.a. Zimmermann et al. 2007; Drews 2008). Um für Nutzer der vorangegangenen Scientific Use Files und dabei insbesondere der IAB-Beschäftigtenstichprobe bestmögliche Kontinuität zu gewährleisten und um die beständige Nachfrage nach einem Regionalfile der Erwerbsverlaufsdaten des IAB zu bedienen, wurde die Stichprobe der Integrierten Arbeitsmarktbiografien in Form eines Regionalfiles aufbereitet und anonymisiert. Die Anonymisierung erfolgte durch ausgewählte Löschungen und durch die Aggregationen von datenschutzrechtlich sensiblen Informationen. Die Anonymisierungsschritte sollen im Folgenden beschrieben werden.

3.3.1 Zusammenfassung von Quellen

Bei der Aufbereitung des SIAB-Regionalfiles aus der SIAB 7508 wurden Angaben zur Quellenherkunft bestimmter Meldungen zusammengefasst. So wurden Informationen aus den Originalquellen der Leistungs-Historiken Grundsicherung im SGB II (LHG und XLHG) sowie zur Arbeitsuche (ASU und XASU) im SIAB-Regionalfile jeweils in einer gemeinsamen Quelle (X)ASU bzw. (X)LHG kombiniert. Dies hat folgende Ursache: die Information darüber, ob eine Person in einer Region von einem zugelassenen kommunalen Träger (zkT) betreut wird, stellt ein Deanonymisierungsrisiko für die beobachteten Person dar, da die Information zum Träger zusammen mit aktuellen, relativ detaillierten Regionalinformationen aus anderen Quellen oder dem Erwerbsverlauf eine gemeindescharfe Lokalisierung ermöglichen kann.

3.3.2 Gültigkeitszeiträume der Quellen

Nach der Aggregation der Datenquellen (vgl. Abschnitt [3.3.1](#)) und gemäß den o.g. datenschutzrechtlichen Anforderungen wurden bei der Aufbereitung des SIAB-Regionalfiles Angaben zum Beginn- bzw. Enddatum ausgewählter Meldungen angepasst. Abbildung 2 veranschaulicht, für welchen Zeitraum die verschiedenen Quellen im SIAB-Regionalfile gefüllt sind.

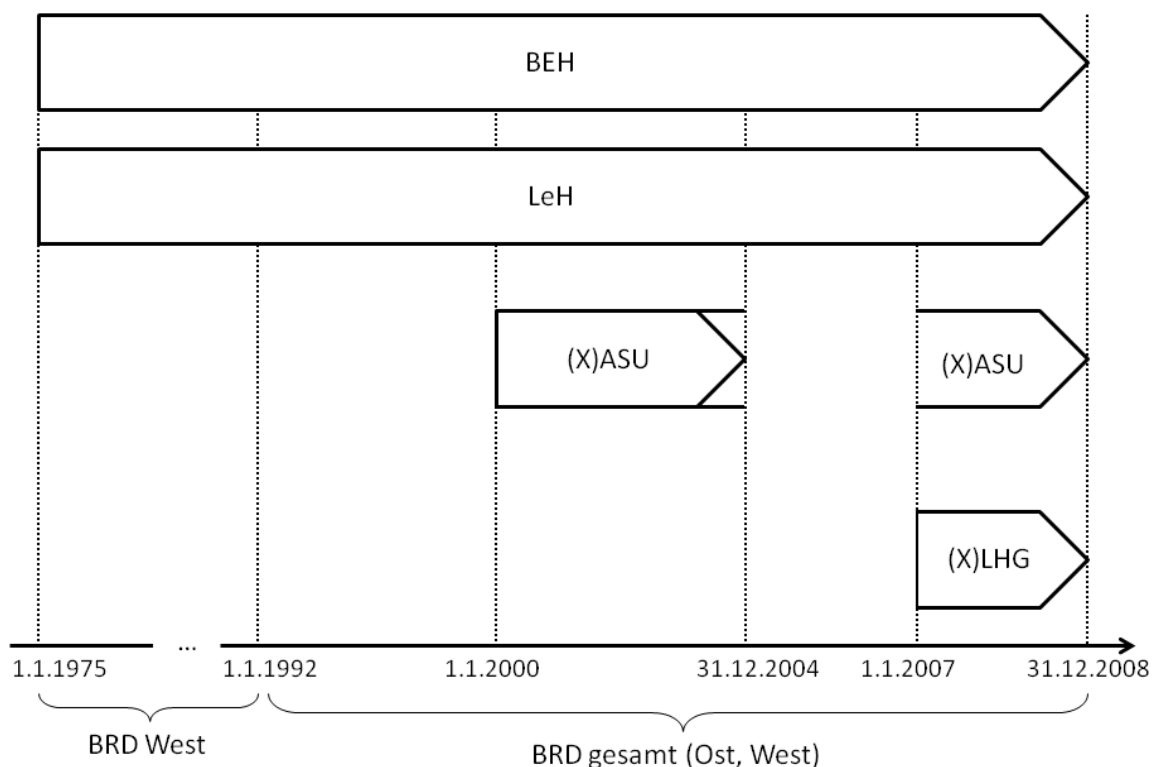


Abbildung 2: Gültigkeitszeiträume der Datenquellen

Die SIAB-R 7508 weist für die BeH und die LeH nur Beobachtungen im Zeitraum zwischen 1.1.1975 und 31.12.2008 aus. Für Meldungen, die Rahmen des Meldeverfahrens zur Sozialversicherung in Ostdeutschland abgegeben wurden, wird erst ab dem Jahr 1992 von deren Vollständigkeit ausgegangen. In Beschäftigungsmeldungen mit gültigen Arbeitsorten in Ostdeutschland vor dem Jahr 1992 wird die Angabe zur Kreisregion auf missing (.z) gesetzt. Somit ist die Unterscheidung, ob eine Person in den alten bzw. neuen Bundesländern beschäftigt ist, im SIAB-Regionalfile erst ab 1992 möglich.

Die (X)ASU ist im SIAB-Regionalfile für den Zeitraum zwischen dem 1.1.2000 und dem 31.12.2004 sowie zwischen dem 1.1.2007 und dem 31.12.2008 gefüllt. Da erst ab dem Jahr 2000 von der Vollständigkeit der ASU Originalquelle ausgegangen wird, werden Meldungen mit einem Enddatum vor dem 1.1.2000 im SIAB-Regionalfile nicht ausgewiesen. Die Löschung von Meldungen in den Jahren 2005 und 2006 ist einerseits bedingt durch die qualitätsrelevanten Untererfassungen von Meldungen der zugelassenen kommunalen Trägern in den IEB Originaldaten (vgl. Dorner et al. 2010). Andererseits ist die Ausweisung wegen des Identifikationsrisikos von Personen, die durch zugelassene kommunale Träger gemeldet werden, in Kombination mit Regionalangaben datenschutzrechtlich problematisch. Die skizzierte Problematik im Zusammenhang mit Meldungen von zugelassenen kommunalen Trägern gilt gleichermaßen für die (X)LHG in den Jahren 2005 und 2006. Entsprechend beinhal-

tet das SIAB-Regionalfile Meldungen der (X)LHG erst ab einem Episodenbeginn zum 1.1.2007.

Beobachtungen, die sowohl mit ihrem Beginn- als auch Enddatum außerhalb dieser quellen-spezifischen Gültigkeitszeiträume liegen, sind nicht mehr in den Daten enthalten. Bei jahres-übergreifenden Meldungen aus den Quellen LeH, (X)ASU sowie (X)LHG, bei denen entwe-der Beginn- oder Enddatum in die quellenspezifischen Gültigkeitszeiträume hinein- bzw. hin-ausreichen, werden Datumsangaben verändert. Dies bedeutet, dass die ursprünglichen Da-tumsangaben entweder zum Beginn- oder Enddatum durch das jeweilige Beginn- bzw. End-datum des quellenspezifischen Gültigkeitszeitraums ersetzt werden.

Eine Besonderheit in den Daten stellen (X)ASU-Sätze dar. Sie können bereits vor dem ge-löschten Zeitraum 2005/2006 beginnen und ohne Unterbrechung bis in das Jahr 2007 rei-chen. Diese Beobachtungen wurden im SIAB-Regionalfile verdoppelt und jeweils das Ende bzw. das Beginndatum an den Gültigkeitszeitraum angepasst, so dass zwei Beobachtungen, ein rechtszensierter und ein linkszensierter Satz, in den Daten enthalten sind.

Folgende fiktive Beispiele in Abbildung 3 sollen die Vorgehensweise bei den Datumskorrek-turen veranschaulichen. Dabei repräsentiert der weiße Balken [1] stets den Originalzeitraum und der grau gefüllte Balken [2] den Zeitraum, der nach der Datumskorrektur im SIAB-Regionalfile enthalten ist.

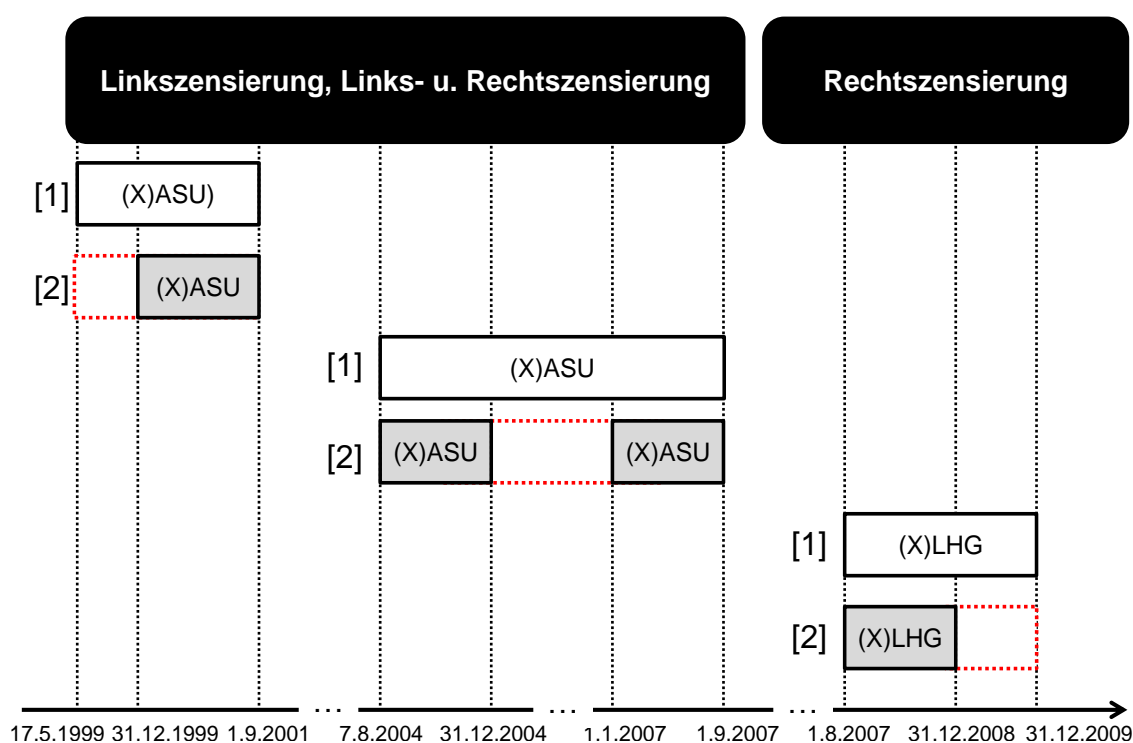


Abbildung 3: Beispiele für Datumskorrekturen jahresübergreifender Meldungen

Im Beispiel zur Linkszensierung einer Arbeitsuchendenmeldung aus der (X)ASU zum 1.1.2000 sowie im Beispiel der Rechtszensierung einer Meldung der (X)LHG (auch LeH oder (X)ASU) zum 31.12.2008 bleibt die Anzahl der Sätze in den Daten gleich. Es werden lediglich die Teile der jahresübergreifenden Meldung gelöscht, die außerhalb des Beobachtungszeitraums liegen, der vom SIAB-Regionalfile abgedeckt wird.

Im mittleren Beispiel, das die Vorgehensweise bei mehrjährigen Beobachtungen anzeigt, welche den in der (X)ASU gelöschten Zeitraum in Jahren 2005 und 2006 überspannen, werden aus der ursprünglichen Meldung zwei zensierte Beobachtungen. Entsprechend erhöht sich die Zahl der Datenzeilen im SIAB-Regionalfile im Vergleich zur Ausgangsdatei.

Eine tabellarische Übersicht zu den durchgeführten Datumsumsetzungen ist dem Anhang (vgl. Abschnitt 7.3, Tabelle A2) zu entnehmen.

3.3.3 Zensierung von Altersangaben

Aus datenschutzrechtlichen Gründen sind nur diejenigen Sätze im Scientific Use File enthalten, bei denen das Alter einer Person zwischen 17 und 62 Jahren liegt. Sätze, die vor dem Alter von 17 enden, bzw. die nach dem Alter von 62 beginnen, sind ausgeschlossen. Bei jahresübergreifenden Sätzen, die vor oder im Jahr des 17. Geburtstags beginnen, jedoch nach diesem Geburtstag enden, bzw. vor oder im Jahr des 62. Geburtstags beginnen und nach dem Jahr des 62. Geburtstags enden, wurde das Originalbeginn- bzw. das Enddatum (begorig bzw. endorig) auf den 1.1. bzw. 31.12. des Jahres umgesetzt, in dem die Person die Altersgrenzen von 17 bzw. 62 erreicht wird. Die Variablen `tage_jung` und `tage_alt` geben jeweils die Anzahl der Tage in (Haupt-)Beschäftigung⁸ vor 17 bzw. nach 62 an.

3.3.4 Aggregation und Löschungen von Variablen

Während bei den Quellen mit Informationen zu Leistungsempfang und Arbeitsuche das Merkmalsspektrum im Vergleich zur schwach anonymisierten Version der SIAB aus Datenschutzgründen stark reduziert wurde, sind die beschäftigungsbezogenen Merkmale aus der BeH weitgehend im SIAB-Regionalfile enthalten geblieben und vergrößert worden.

Die im SIAB-Regionalfile verbleibenden Merkmale bilden die Grundlage für die Anwendung weiterer Anonymisierungsschritte in Form von Informationsaggregation. Müller et al. (1991, S. 444 f.) schlagen für die Prüfung und die Erstellung eines faktisch anonymen Regionalfiles die drei folgenden Regeln vor:

⁸ Die Berechnung der Beschäftigungszeiten berücksichtigt folgende Restriktionen: `quelle_gr==1 & level2==0` (Stata Syntax)

1. *Durch Kombination von Regionalklassifikationen darf keine in den Daten ausgewiesene Regionaleinheit ermittelbar sein, die eine Einwohnerzahl von weniger als 100.000 Personen aufweist. (vgl. Müller et al. 1991, S. 444)*

Die Umsetzung dieser Vorgabe erfolgte anhand der Prüfung der Einwohnerzahlen der 413 kreisfreien Städte und Landkreise zum 31.12.2008⁹. Die im SIAB-Regionalfile enthaltene Regionalisierung fasst die Kreisangaben in den Ausgangsdaten so zusammen, dass jede der entstehenden Kreisregionen zum Stichtag eine Einwohnerzahl von mindestens 100.000 Personen aufweist. Diese Regionalisierung im SIAB-Regionalfile lehnt sich an die Vorgehensweise der Bildung von Mikrozensuskreisregionen (MZKR) an (vgl. Christians/Wirth 2009, S.22). Die Aggregation führt zu insgesamt 333 Kreisregionen, die im Merkmal Region des Arbeitsortes (ao_region) ausgewiesen werden (Gebietsstand 31.12.2008). Bei der Bildung der Kreisregionen wurden wie bei der Erstellung der MZKR übergeordnete Raumeinheiten wie Raumordnungsregionen und Bundesländer berücksichtigt (vgl. Abschnitt 7.5, Abbildung A1/Tabelle A10).

Ergänzend wurde das Merkmal „Pendlerstatus“ (pendler) generiert, welches anzeigt, ob sich der aktuelle Wohnort vom aktuellen oder letzten ausgewiesenen Arbeitsort unterscheidet.

2. *Potenziell zur Deanonymisierung geeignete Merkmale müssen so stark aggregiert werden, dass keine Merkmalsausprägungen mehr ausgewiesen werden, die*
 - a. *in der Grundgesamtheit nicht wenigstens 50.000 Personen umfassen.*
 - b. *pro übermittelter Regionaleinheit nicht mindestens drei Personen im Datensatz enthalten; Merkmalsausprägungen, die im Datensatz nur einen oder zwei Fälle enthalten, dürfen ausschließlich in einer stärker aggregierten Form ausgewiesen werden. (vgl. Müller et al. 1991, S. 444 f.)*
3. *Alle übrigen im Regionalfile enthaltenen Merkmale müssen – falls unter Datenschutzgesichtspunkten erforderlich – so weit aggregiert werden, dass jede ausgewiesene Merkmalsausprägung in der Grundgesamtheit mindestens 5.000 Personen umfasst. (vgl. Müller et al. 1991, S. 445)*

Auf Basis dieser genannten Kriterien wurden bei der Erstellung des SIAB-Regionalfiles weitere Merkmale aggregiert. Grundlage dazu bilden Häufigkeitsauszählungen von Personen

⁹ Datenbasis: Bevölkerungsfortschreibung des Statistischen Bundesamtes zum 31.12.2008 für die kreisfreien Städte und Kreise der BRD

zum 30.06. eines jeden Jahres. Merkmalsausprägungen von Variablen mit hohem Deanonymisierungsrisiko wurden zum einen aggregiert, wenn deren univariate Randverteilung zum 30.06. des jeweiligen Jahres eine Mindestanzahl an Personen nicht umfasst (2a). Zum anderen wurden sie aggregiert, wenn in den Zellen der Kreuztabellen zwischen dem jeweiligen Merkmal und dem Merkmal Arbeitsort Region nicht mindestens 3 Personen enthalten sind (2b). Bei der Prüfung der Wirtschaftszweigangaben wurde nicht die Anzahl der Personen, sondern die der Betriebe zugrunde gelegt, da es sich um ein Betriebsmerkmal handelt. Alle weiteren Merkmale werden über den gesamten Beobachtungszeitraum der Daten ausgezählt und falls erforderlich auf Basis von (3) aggregiert, wenn zu einem Zeitpunkt in den Datenzellen der SIAB 7508 die o.g. Mindestfallzahlen unterschritten werden.

Liegen zu einer Person am Stichtag parallele Informationen aus unterschiedlichen Quellen vor, z.B. Arbeitsuche und zeitgleicher Leistungsempfang, so wurde die Person in einem Merkmal (z.B. Alter oder Staatsangehörigkeit), das in den beiden Quellen gefüllt ist, bei den quellenübergreifenden Auszählungen nur einmal berücksichtigt.

Bei folgenden Merkmalen wurden Aggregationen durchgeführt: Staatsangehörigkeit, Beruf – ausgeübte Tätigkeit, Klassifikation der Wirtschaftszweige (WS 73 und WZ 2003), Abgabegrund/Beendigungsgrund/Abmeldegrund, Erwerbsstatus. Für diese Merkmale wurden die Jahre 1975 bis einschließlich 1984 nicht zur Beurteilung des für die faktische Anonymität der Merkmale benötigten Aggregationsniveaus herangezogen. Aufgrund des strukturellen Wandels auf dem Arbeitsmarkt hätte eine strikte Auslegung der o.g. Anonymisierungsregeln in diesem Zeitraum ein sehr hohes Aggregationsniveau und damit eine starke Einschränkung des Analysepotenzials der Daten für bestimmte Personengruppen zur Folge gehabt. Das Tagesentgelt sowie die Leistungssätze wurden auf volle Euro gerundet.

3.3.5 Auswirkungen der Anonymisierung

Die Informationen zu Beschäftigung und Leistungsempfang nach SGB III im SIAB Regionalfile sind nur in geringem Maße von Löschungen auf Satz- bzw. Personenebene betroffen. Anonymisierungsbedingte Löschungen einzelner Sätze im Zuge der Anpassung der Beobachtungszeiträume für die Quellen (X)ASU und (X)LHG resultierten im Wegfall von einigen Personen aus dem SIAB-Regionalfile. Durch die o.g. Aufbereitungsschritte wurden bezogen auf die SIAB 7508 insgesamt 143.561 Personenkonten gelöscht. Das SIAB-Regionalfile beinhaltet damit noch 91,3 Prozent der in der Ausgangsdatei SIAB 7508 enthaltenen Personen (vgl. Abschnitt 7.3, Tabelle A3).

Einige personen- und betriebsbezogene Merkmale wurden über alle Quellen hinweg entweder vergrößert oder im Zuge der Anonymisierung aus den Daten gelöscht. Dabei wurde bei den Quellen mit Informationen zu Leistungsempfang und Arbeitssuche das Merkmalspektrum deutlich reduziert, während die beschäftigungsbezogenen Merkmale weitgehend erhalten geblieben und lediglich vergrößert worden sind. Mit dem SIAB-Regionalfile ist somit die Kontinuität für beschäftigungsbezogene Analysen im Vergleich zur IAB Beschäftigtenstichprobe (insb. IABS-Regionalfile 1975-2004) im hohen Maße gegeben.

Detailliertere Informationen zu anonymisierungsbedingten Löschungen und Aggregationen auf Merkmalsebene im Vergleich zur schwach anonymisierten Version der SIAB 7508 sowie im Vergleich zu verwandten Datenprodukten des FDZ finden Sie im tabellarischen Anhang des Datenreports.

3.4 Definition fehlender, ungültiger sowie anonymisierter Werte

In der SIAB-R 7508 werden fehlende bzw. durch die Anonymisierung nicht mehr als gültig ausgewiesene Werte wie folgt kodiert:

Bezeichnung	Wert	Beschreibung
Keine (valide) Angabe, anonymisierte Angabe	.z	Nicht systematisch fehlende Ausprägungen eines Merkmals, d.h. das Merkmal ist prinzipiell für die Quelle gefüllt, jedoch ist für die betrachtete Ausprägung keine Angabe vorhanden bzw. keine Zuordnung möglich. Außerdem: Aus datenschutzrechtlichen Gründen anonymisierte Werte. Eine Unterscheidung zwischen aus anonymisierten Angaben und ungültigen Angaben in den Ausgangsdaten ist nicht möglich.
Systematisch nicht gefüllt	.n	Ein Merkmal ist grundsätzlich für eine Quelle (dunkelgraue Felder der Merkmalsübersicht im Abschnitt 1.5) oder für einen gewissen Zeitraum nicht gefüllt.

Tabelle 4: Kodierung fehlender und anonymisierter Werte

3.5 Episodensplitting

Die Daten des SIAB-Regionalfiles liegen „gesplittet“ vor. Falls sich innerhalb eines Kontos Sätze aus den Quellen überlappen, werden diese durch künstliche Sätze mit neuen Datumsangaben derart ersetzt, dass vollständig parallele bzw. überschneidungsfreie Zeiträume entstehen (Episodensplitting). Die Anzahl der in den Daten enthaltenen Sätze erhöht sich dadurch.

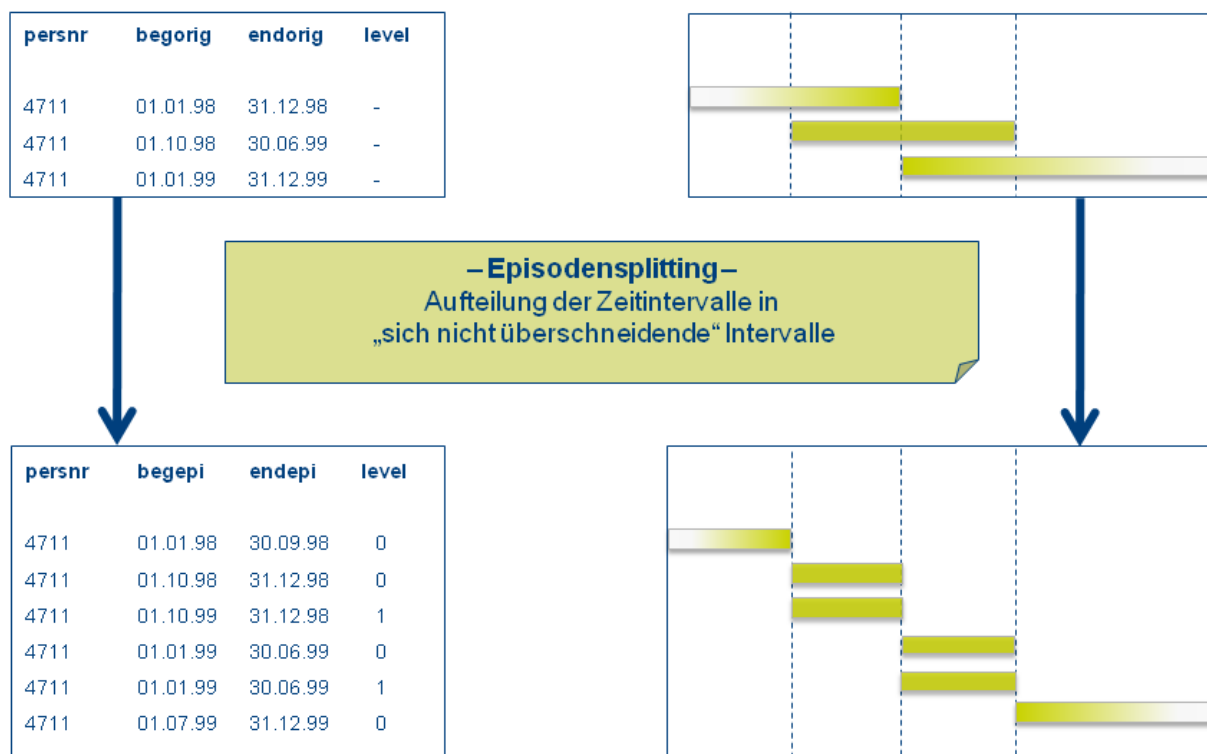


Abbildung 4: Episodensplitting in den IEB und SIAB Daten

Durch das Episodensplitting werden Sätze verdoppelt und der Gültigkeitszeitraum geändert. Die ursprünglichen Datumsvariablen Beginn und Ende des Originalsatzes (begorig und endorig) bilden hier den Ausgangspunkt für das Episodensplitting. Die Variablen Beginn und Ende der Episode (begepi und endepi) kennzeichnen Anfang und Ende der gesplitteten Episoden. Ob durch das Episodensplitting Originalsätze vervielfältigt wurden, lässt sich anhand zusätzlich generierter technischer Merkmale im Datensatz (level1, level2) feststellen. Das Episodensplitting kann in den Daten rückgängig gemacht werden, indem alle Beobachtungen behalten werden, in denen das Beginndatum der Episode dem Beginndatum des Originalsatzes entspricht (Stata Syntax: `keep begorig == begepi`).

4 Datenqualität und Probleme

4.1 Beschäftigten-Historik (BeH)

- Aufgrund der Einführung des Meldeverfahrens in den neuen Bundesländern kann bei BeH-Sätzen erst ab 1993 von einer hinreichenden Vollständigkeit der Ost-Meldungen ausgegangen werden. Der Anstieg der Zahl der BeH-Sätze ab 1999 folgt aus der Einführung der Meldepflicht für geringfügig Beschäftigte ab dem 1. April 1999.
- Innerhalb des Meldeverfahrens ist eine gewisse Zeitverzögerung nicht vermeidbar. Zwar müssen Änderungen von Beschäftigungsverhältnissen sofort gemeldet und bestehende Beschäftigungsverhältnisse jährlich bis zum April des folgenden Jahres bestätigt werden, faktisch treffen Meldungen jedoch zum Teil um Jahre verspätet ein. Die Historik-Datei des IAB wird aber nicht kontinuierlich fortgeschrieben, sondern in bestimmten Abständen aktualisiert. Dies geschieht unter Verwendung von Dateien mit den Beschäftigten-Meldungen jeweils eines Jahres, die 36, 18, 12 oder 6 Monate nach Ablauf des Berichtsjahres eingetroffen sind (frühestens im Juli 2009 z. B. kann die 18-Monats-Datei für das Jahr 2007 erstellt werden). Mehr als 3 Jahre verspätete Meldungen werden im IAB nicht berücksichtigt, sodass eine 36-Monats-Datei per definitionem einen Füllgrad von 100 % aufweist.
- In der den SIAB-Daten zugrundeliegenden Version der IEB liegt der Füllgrad der BeH-Meldungen im Jahr 2006 letztmalig bei 100 %. Für das Jahr 2007 wurde eine 18-Monatsdateien verwendet, die Meldungen für das Jahr 2008 stammen aus einer 12-Monats-Datei. Es ist daher davon auszugehen, dass die BeH-Meldungen für die Jahre 2007 und insbesondere 2008 in der SIAB leicht untererfasst sind. Die Auswertbarkeit der Daten auf Personenebene sollte dadurch aber nicht eingeschränkt sein. Allerdings häufen sich die fehlenden Meldungen in wenigen Betrieben.
- Im Jahr 1984 erfolgte eine Änderung des Meldeverfahrens. Einmalig gezahltes Bruttoarbeitsentgelt wurde ab diesem Zeitpunkt als Teil des beitragspflichtigen Jahresentgelts gemeldet, was zu einer Erhöhung des durchschnittlichen Tagesentgelts führt. Insbesondere steigt ab diesem Jahr der Anteil der Entgelte oberhalb der Beitragsbemessungsgrenze der deutschen Rentenversicherung deutlich an¹⁰.
- Für einzelne Geburtsjahrgänge zeigt sich im Jahr 1975 eine geringere Fallzahl als erwartet. Das Jahr 1975 ist somit für Auswertungen problematisch¹¹.
- Bedingt durch die Einführung des neuen Meldeverfahrens 1999 sind nur die Personengruppen 101, 102, 103, 105, 106, 112, 118, 119, 120, 140, 141, 142, 143, 201,

¹⁰ Vgl. Bender et al. 1996.

¹¹ Vgl. Bender et al. 1996.

203 und 205 enthalten, sowie die beiden Gruppen 109 und 209, die geringfügig Beschäftigte kennzeichnen. Nicht enthalten sind unter anderem kurzfristig Beschäftigte (Personengruppen 110, 202 und 210). Im SIAB-Regionalfiler sind o.g. Personenkreise alle in aggregierter Form enthalten. Ausnahme: Vor 1999 sind zusätzlich Meldungen erfasst, die keine (gültige) Angabe zur Personengruppe aufweisen, was insbesondere für das Jahr 1991 von Bedeutung ist, weil dort viele Beschäftigte in Ostdeutschland so gemeldet wurden.

4.2 Leistungsempfänger-Historik (LeH)

- Die Leistungsbezugsdaten wurden früher auf Magnetbändern gesichert. Aufgrund eines Schadens an den Magnetbändern sind die Leistungsbezugsdaten bis einschließlich des Jahres 1980 nur teilweise enthalten. Daher ist auch in den SIAB-Daten davon auszugehen, dass Informationen zum Leistungsbezug im genannten Zeitraum nur zum Teil enthalten sind. Dies ist bei Analysen der LeH entsprechend zu berücksichtigen.

4.3 Leistungs-Historiken Grundsicherung (LHG/XLHG)

- Hinsichtlich der Vollständigkeit der Fallzahlen bzw. Leistungsbiografien aus den Quellen LHG und XLHG, bestehen in den IEB ab 2007 noch punktuelle Lücken bei einzelnen Trägern. Es ist davon auszugehen, dass sich diese Lücken auch in den SIAB-Daten fortsetzen. Darüber hinaus ist die Qualität der Meldungen in hohem Maß vom meldenden Träger abhängig, was besonders bei regionalen Analysen zu beachten ist.
- In den SIAB-Daten ist der Füllgrad des Merkmals „SGB-Einstellungsgrund“ („grund_gr“, vgl. 5.4.1) relativ gering. Hier ist jedoch zu beachten, dass der „SGB-Einstellungsgrund“ ein rein personenbezogenes Merkmal ist. Wird daher ein Satz einer Person im ALG-II-Bezug aufgrund einer Änderung der Struktur seiner Bedarfsgemeinschaft gesplittet, weist der „SGB-Einstellungsgrund“ (grund_gr) die Ausprägung missing (.z) auf.
- Beginn und Ende eines Sanktionssachverhalts führen zwar zu einem Splitting der Sätze, jedoch kann anhand der Daten weder die Dauer, noch die Art oder der Zeitpunkt der Verhängung bzw. des Beginns der Sanktion festgestellt werden.
- In der LHG als auch im höheren Maß in der XLHG ist zu beobachten, dass es sowohl in den IEB überdurchschnittlich viele 15-Jährige und in geringerem Maß 16- und 17-Jährige mit Hinweisen auf fehlende Erwerbsfähigkeit gibt. 15- bzw. 16-jährige Hilfe-

bedürftige der zugelassenen kommunalen Träger sind deshalb möglicherweise schlechter in den IEB repräsentiert, da „nicht Erwerbsfähige“ in die IEB nicht übernommen werden.

4.4 Arbeitssuchenden-Historiken (ASU/XASU)

ASU

- Die registrierten Zeiten der Arbeitsuche mit einem Meldeende ab dem Jahr 2000 für die einzelnen Personen werden als vollständig angesehen.

XASU

- Im Scientific Use File ist ab 2007 von einer Vollständigkeit der XASU-Meldungen von den 69 zugelassenen kommunalen Trägern in der IEB auszugehen.
- Wegen der fehlenden zentralen Zuordnung von Identifikationsnummern sind Brüche oder unplausible Arbeitsuche-Biografien in den IEB und damit auch in den SIAB-Daten möglich.
- Einige Merkmale im Scientific Use File weisen für die XASU zum Teil nur einen sehr geringen Füllgrad auf. Hiervon betroffen sind besonders die Merkmale Staatsangehörigkeit, Abgabegrund und Erwerbsstatus. Die Füllgrade dieser Merkmale verbessern sich zwar im Zeitablauf, bleiben aber dennoch zum Teil unbefriedigend. Generell ist zu beachten, dass die Füllgrade zum Teil auch trägerspezifisch variieren.
- Weiterhin ist unklar, ob die Träger - wie vorgesehen - nur Zeiten melden, in denen sie selbst die Arbeitssuchenden betreuen, oder ob vorausgehende Zeiten der Betreuung durch die BA nach dem SGB III als durchgängige Episoden gemeldet werden, sofern der Arbeitsuche-Status nicht wechselt.
- Bei etlichen Trägern ist in den IEB durchgängig oder zeitweise der Anteil der gemeldeten Arbeitssuchenden an den gemeldeten erwerbsfähigen Hilfebedürftigen unplausibel hoch. Ein Grund dafür könnte eine möglicherweise fehlerhafte Bestimmung des Status „Nichtarbeitslos arbeitssuchend“ bei diesen Trägern sein.

5 Merkmalsbeschreibungen

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in gesonderten Dateien unter <http://fdz.iab.de>.

5.1 Identifikatoren

5.1.1 Systemfreie Personennummer (persnr)

Variablenlabel	Systemfreie Personennummer
Variablenname	persnr
Kategorie	Identifikatoren
Herkunft	BeH, LeH, (X)LHG, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Die systemfreie Personennummer zeigt an, welche Sätze zur gleichen Person gehören. Systemfrei bedeutet, dass von dieser Personennummer nicht auf Eigenschaften der Person oder Original-Identifikatoren geschlossen werden kann.</p> <p>Da es keinen einheitlichen BA-weiten Personen-Identifikator gibt, ist die Zuordnung von (BA-)Daten aus unterschiedlichen Datenquellen (z. B. Beschäftigung und Leistung) zu Personen nicht immer eindeutig. Die Bildung des quellenübergreifenden Personen-Identifikators basiert auf einer von der BA entwickelten Heuristik.</p> <p>Die systemfreie Personennummer aus der SIAB 7508 wurde im SIAB-Regionalfiler durch eine systemfreie, zufällig sortierte Personennummer ersetzt. Der Wertebereich dieser neu generierten Personennummer erstreckt sich von 1 bis 1.515.463.</p>

5.1.2 Betriebsnummernzähler (bnn)

Variablenlabel	Betriebsnummernzähler
Variablenname	bnn
Kategorie	Identifikatoren
Herkunft	BeH,
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p>Der Betriebsnummernzähler nummeriert die Betriebe im Erwerbsleben einer Person im Erwerbsleben aufsteigend.</p> <p>Beispiel: Der erste Betrieb, in dem eine Person beschäftigt war, erhält den Wert 1. Wechselt die Person zu einem anderen Betrieb, erhält dieser den Wert 2. Für nachfolgende Betriebe erhöht sich der Wert um jeweils 1. Wechselt die Person zu einem Betrieb, in welchem sie zu einem früheren Zeitpunkt bereits beschäftigt war, so erhält dieser Betrieb den Wert, der für die erste Beschäftigung galt (z.B. 2). Wechselt eine Person nach genau einem Wechsel zum ersten Betrieb zurück, ergäbe sich daher im Zeitverlauf für die Variable bnn die Abfolge 1-2-1. Fehlende Betriebsnummern in den Originaldaten wurden auch im SUF mit missing (.z) gekennzeichnet.</p>
--------------------	--

5.2 Gültigkeitszeitraum

5.2.1 Beginndatum Originalsatz (begorig)

Variablenlabel	Beginndatum Originalsatz
Variablenname	begorig
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH, (X)LHG, (X)ASU
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Beginndatum des Originalsatzes entspricht dem ursprünglichen Beginndatum der Meldung. Durch die o.g. Datumskorrekturen der Variable begorig kann diese im SIAB-Regionalfile von der Angabe in SIAB abweichen. begorig kann vom Beginndatum der Episoden (begepi) abweichen. Grund dafür ist das so genannte Episodensplitting (vgl. Abschnitt 3.7).</p> <p>In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe von Jahresmeldungen).</p>

5.2.2 Enddatum Originalsatz (endorig)

Variablenlabel	Enddatum Originalsatz
Variablenname	endorig
Kategorie	Gültigkeitszeitraum
Herkunft	BeH, LeH, (X)LHG, (X)ASU
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	<p>Das Enddatum des Originalsatzes entspricht dem ursprünglichen Enddatum der Meldung. Durch die o.g. Datumskorrekturen der Variable endorig kann diese im SIAB-Regionalfile von der Angabe in SIAB abweichen. endorig kann vom Enddatum der betreffenden Datenzeile, dem so genannten Episoden-Enddatum, abweichen. Grund dafür ist das so genannte Episodensplitting (vgl. Abschnitt 3.7).</p> <p>In BeH-Meldungen sind Beginn- und Endjahr wegen der Regeln des Meldeverfahrens immer identisch (Pflicht des Arbeitgebers zur Abgabe</p>

	von Jahresmeldungen).
--	-----------------------

5.2.3 Beginndatum der Episode (begepi)

Variablenlabel	Beginndatum der Episode
Variablenname	begepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH, (X)LHG,(X)ASU
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Beginndatum der Episode. Mit dem Begriff „Episode“ wird der Zeitraum eines Satzes bezeichnet. Dieser Zeitraum kann durch Rechts- bzw. Linkszensierung einzelner Episoden (siehe Kapitel 3) und/oder durch Episodensplitting vom Originalzeitraum der ursprünglichen Meldung abweichen (vgl. Abschnitt 3.7).

5.2.4 Enddatum der Episode (endepi)

Variablenlabel	Enddatum der Episode
Variablenname	endepi
Kategorie	Generierter Gültigkeitszeitraum
Herkunft	BeH, LeH, (X)LHG,(X)ASU
Datentyp	Datum
Hierarchie	Keine
Detailbeschreibung	Enddatum der Episode. Mit dem Begriff „Episode“ wird der Zeitraum eines Satzes bezeichnet. Dieser Zeitraum kann durch Rechts- bzw. Linkszensierung einzelner Episoden (siehe Kapitel 3) und/oder durch Episodensplitting vom Originalzeitraum der ursprünglichen Meldung abweichen (vgl. Abschnitt 3.7).

5.3 Informationen zur Person

5.3.1 Geschlecht (frau)

Variablenlabel	Geschlecht
Variablenname	frau
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, (X)LHG,(X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Geschlechter-Dummy (0 - Mann, 1 - Frau). Die Geschlechtsangabe ist pro Konto konstant.
Qualitätshinweis	In den Originaldaten kann es vorkommen, dass das Geschlecht wechselt. Dies wird bei der Datenaufbereitung bereinigt. Hierbei kommt Angaben aus Datenquellen der Leistungsgewährung (LeH, (X)LHG) die höchste Priorität zu.

5.3.2 Geburtsjahr (gebjahr)

Variablenlabel	Geburtsjahr
Variablenname	gebjahr
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, (X)LHG,(X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Geburtsjahr ist pro Konto konstant. Sätze mit einem Alter kleiner als 17 bzw. einem Alter größer als 62 Jahre wurden gelöscht (siehe Abschnitt 3.3.3).
Qualitätshinweis	In den Originaldaten kann es vorkommen, dass das Geburtsjahr zwischen den Quellen wechselt. Dies wird bei der Datenaufbereitung bereinigt. Hierbei kommt der Angabe aus der Sozialversicherungsnummer die höchste Priorität zu.

5.3.3 Staatsangehörigkeit (deutsch)

Variablenlabel	Staatsangehörigkeit deutsch
Variablenname	deutsch
Kategorie	Persönliche Merkmale
Herkunft	BeH, LeH, (X)LHG, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal unterscheidet für Personen, deren Arbeitsort in Westdeutschland liegt, zwischen deutscher und nicht-deutscher Staatsangehörigkeit. Für Personen, die in ihrem Erwerbsverlauf mindestens einmal in Ostdeutschland beschäftigt waren, ist keine Angabe zur Staatsangehörigkeit enthalten. Ebenfalls ausgewiesen wird die Staatsangehörigkeit bei Personen, in denen bei keiner Meldung ein gültiger Arbeitsort vorliegt und somit die Zuordnung zu Ost- bzw. Westdeutschland nicht möglich ist.

5.3.4 Ausbildung (bild)

Variablenlabel	Ausbildung
Variablenname	bild
Kategorie	Persönliche Merkmale
Herkunft	BeH, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p>Zu beachten ist, dass das Merkmal je nach Quelle eine unterschiedliche Bedeutung hat</p> <p>1) BeH</p> <p>Das Merkmal enthält für BeH-Sätze die von den Arbeitgebern im Rahmen des Meldeverfahrens anzugebende Ausbildung. Bei dem Signierschlüssel „Ausbildung“ handelt es sich um einen kombinierten Schlüssel, durch den sowohl die erreichte Schulbildung als auch die abgeschlossene Berufsausbildung (jeweils höchster Abschluss) der/des Beschäftigten erhoben werden soll. Dabei ist grundsätzlich zunächst die Schulbildung festzustellen, danach die Berufsausbildung. Lediglich bei den Schlüsselpositionen 5 (Abschluss einer Fachhochschule) und 6 (Hochschul- / Universitätsabschluss) wird auf die Feststellung einer sonstigen Berufsausbildung verzichtet (vgl. BA 2005, S. IX.). Folgende Ausprägungen kommen vor:</p> <ol style="list-style-type: none"> 1 Volks-/Hauptschule/mittlere Reife oder gleichwertige Schulbildung, ohne abgeschlossene Berufsausbildung 2 Volks-/Hauptschule/mittlere Reife oder gleichwertige Schulbildung, mit abgeschlossener Berufsausbildung 3 Mit Abitur, ohne abgeschlossene Berufsausbildung 4 Mit Abitur und abgeschlossener Berufsausbildung 5 Fachhochschulabschluss 6 Hochschulabschluss <p>Aus den Daten kann nicht geschlossen werden, ob eine abgeschlossene Schulausbildung vorliegt oder nicht, beide Fälle können in den Ausprägungen „1“ und „2“ vorkommen.</p> <p>„Zusammen mit einem Betriebswechsel treten häufig auch Änderungen im Ausbildungsstatus auf. Das liegt daran, dass beim neuen Betrieb die Meldedaten neu zusammenstellt werden. Wenn ein Beschäftigter z.B. durch berufsbegleitende Weiterbildung einen höheren Abschluss erlangt hat, dann wird dieser Statuswechsel wahrscheinlich erst von einem neuen Beschäftigungsbetrieb erfasst. Es ist in der Regel davon auszugehen, dass bei längeren Beschäftigungszeiten innerhalb eines Betriebs die einmal erhobenen personenbezogenen Daten fortgeschrieben werden.“ (Meinken/Koch 2004, S. 63)</p> <p>Eine Methode, fehlende Werte oder inkonsistente zeitliche Verläufe der Bildungsangaben in der Vorgängerstichprobe IABS zu bereinigen, findet sich in Fitzenberger et al. (2006) sowie in Drews (2006). Hierfür wird aber nur die Quelle BeH herangezogen, die in der IABS die einzige Datenquelle mit Informationen zur Bildung war.</p> <p>2) (X)ASU</p> <p>Für (X)ASU-Sätze ist die abgeschlossene Berufsausbildung zum Beginn der Arbeitsuche enthalten. Folgende Ausprägungen kommen vor:</p> <ol style="list-style-type: none"> 21 ohne abgeschlossene Berufsausbildung 22 betriebliche Ausbildung 23 außerbetriebliche Ausbildung 24 Berufsfachschule 25 Fachschule 26 Fachhochschule
--------------------	--

	27 Universität
Qualitätshinweis	„Für bestimmte Teilgruppen gibt es einen hohen Anteil fehlender Angaben, weil das Merkmal versicherungsrechtlich keine besondere Bedeutung hat (das betrifft z.B. die geringfügig Beschäftigten).“ (Meinken/Koch 2004, S. 63).

5.3.5 Schulausbildung (schbild)

Variablenlabel	Schulausbildung
Variablenname	Schbild
Kategorie	Persönliche Merkmale
Herkunft	(X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Das Merkmal enthält den Schulabschluss zu Beginn der Arbeitslosigkeit bzw. Arbeitsuche.
Qualitätshinweis	Der Füllgrad in der XASU steigt sich zwar kontinuierlich bis 2008, bleibt aber problematisch. Die Quellen ASU und XASU können allerdings im faktisch anonymisierten Datensatz nicht unterschieden werden.

5.4 Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche

5.4.1 Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund/ Abmeldegrund (grund_gr)

Variablenlabel	Abgabegrund /Beendigungsgrund/ SGB-II-Einstellungsgrund /Abmeldegrund, vergrößert
Variablenname	grund_gr
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, (X)LHG, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>1) BeH</p> <p>In BeH-Meldungen kennzeichnet der Abgabegrund den Anlass, aus dem der Arbeitgeber den Sozialversicherungsträgern die betreffende Beschäftigungsmeldung übermittelt hat. Allerdings treten in der IEB und damit auch in der SIAB nicht alle im Rahmen des Meldeverfahrens möglichen Abgabegründe auf. So beinhaltet die IEB nur Meldungen mit Entgeltangaben (also Jahres-, Unterbrechungs- und Abmeldungen), während Anmeldungen mit fehlenden Entgeltinformationen nicht enthalten sind. Ein Informationsverlust ist damit jedoch nicht verbunden, da die Angaben aus einer Anmeldung mit der darauf folgenden Jahres-, Unterbrechungs- oder Abmeldung erneut übermittelt werden. Die Abgabegründe für Beschäftigungsmeldungen sind in der SIAB</p>

	<p>nach den Regeln des seit 1.1.1999 geltenden Meldeverfahrens (gemäß DEÜV) kodiert.</p> <p>2) LeH</p> <p>In LeH-Meldungen gibt das Merkmal den Grund für das Ende des Bezugs von Arbeitslosengeld, -hilfe oder Unterhaltsgeld an. Über die Gründe für den Beginn des Leistungsbezugs finden sich in der LeH und damit auch in der SIAB hingegen keine Informationen, da die LeH mit den Meldungen der Arbeitsagenturen an die Krankenkassen über abgeschlossene Leistungsbezugsdauern gefüllt wird.</p> <p>3) (X)LHG</p> <p>Bei (X)LHG Datensätzen ist der SGB-II-Einstellungsgrund enthalten und gibt den Grund für die Beendigung von laufenden Leistungen an. Der SGBII-Beendigungsgrund bezieht sich auf die Person, nicht aber auf die Bedarfsgemeinschaft. Das Merkmal ist genau zum Satz-Ende gültig. Bei einem Wechsel seines Wertes wird der Datensatz gesplittet. Es ist zu beachten, dass der SGB-II-Beendigungs-/Einstellungsgrund ein personenbezogenes Merkmal ist. Wird bei einem anderen als dem betrachteten Mitglied der Bedarfsgemeinschaft der SGBII-Leistungsbezug beendet, werden alle Sätze der Mitglieder in einer Bedarfsgemeinschaft an diesem Datum gesplittet, der SGBII-Beendigungs-/Einstellungsgrund ist jedoch nur für die betroffene Person gefüllt.</p> <p>4) (X)ASU</p> <p>Bei (X)ASU-Sätzen ist der Abmelde- bzw. Abgangsgrund enthalten. Hierbei ist zu beachten, dass ab dem 26.04.2003 eine Reduktion der Merkmalsausprägungen erfolgte. Die alten Werte sind, soweit es möglich war, auf die neuen Werte umgeschlüsselt.</p> <p>Im Zuge der Anonymisierung der SIAB wurden Abgabegründe innerhalb der vier im SIAB-Regionalfile enthaltenen Quellen zusammengefasst.</p>																																			
Qualitätshinweis	<p>Anteil valider Werte (Füllgrad) des Abgabegrundes in den Quellen (X)LHG und (X)ASU</p> <table><tr><td colspan="2"></td><td colspan="5">Jahr</td></tr><tr><td>quelle_gr</td><td></td><td>2000</td><td>2001</td><td>2002</td><td>2003</td><td>2004</td></tr><tr><td>BeH</td><td> </td><td>1.000</td><td>1.000</td><td>1.000</td><td>1.000</td><td>1.000</td></tr><tr><td>LeH</td><td> </td><td>1.000</td><td>1.000</td><td>1.000</td><td>0.680</td><td>0.695</td></tr><tr><td>(X) ASU</td><td> </td><td>0.794</td><td>0.775</td><td>0.784</td><td>0.769</td><td>0.753</td></tr></table>			Jahr					quelle_gr		2000	2001	2002	2003	2004	BeH		1.000	1.000	1.000	1.000	1.000	LeH		1.000	1.000	1.000	0.680	0.695	(X) ASU		0.794	0.775	0.784	0.769	0.753
		Jahr																																		
quelle_gr		2000	2001	2002	2003	2004																														
BeH		1.000	1.000	1.000	1.000	1.000																														
LeH		1.000	1.000	1.000	0.680	0.695																														
(X) ASU		0.794	0.775	0.784	0.769	0.753																														

-----+-----				
		Jahr		
quelle_gr		2005	2006	2007 2008
-----+-----				
BeH		1.000	1.000	1.000 1.000
LeH		0.969	1.000	1.000 1.000
(X) ASU				0.913 0.862
(X) LHG				0.156 0.131
-----+-----				

5.4.2 Tagesentgelt, täglicher Leistungssatz (tentgelt_gr)

Variablenlabel	Tagesentgelt, täglicher Leistungssatz
Variablenname	tentgelt_gr
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>1) BeH</p> <p>In BeH-Sätzen weist das Merkmal jeweils das Brutto Tagesentgelt des Beschäftigten aus. Es wurde berechnet aus dem vom Arbeitgeber gemeldeten Zeitraumentgelt sowie der Dauer des Zeitraums der ungesplitteten Originalmeldung der IEB in Kalendertagen. Das Tagesentgelt ist für den gesamten Zeitraum einheitlich in der Währung Euro angegeben. Im SIAB-Regionalfiler wird das Tagesentgelt aus datenschutzrechtlichen Gründen nur gerundet in vollen Euro Werten ausgewiesen. Entgelte unter 0,5 Euro wurden dabei auf 0 Euro abgerundet.</p> <p>Bis 1998 meldeten Arbeitgeber grundsätzlich nur das sozialversicherungspflichtige Arbeitsentgelt. Entgelte, die unterhalb der Geringfügigkeitsgrenze lagen, wurden nicht gemeldet.</p> <p>Entgelte, welche die Beitragsbemessungsgrenze in der gesetzlichen Rentenversicherung überschritten, werden nur bis zu dieser Grenze gemeldet. Hierbei ist zu beachten, dass es in der gesetzlichen Rentenversicherung zwei Beitragsbemessungsgrenzen gibt: die Grenze der Arbeiterrenten- und Angestelltenversicherung und die Grenze der Knappschaftlichen Rentenversicherung, die meist höher liegt. Eine Unterscheidung welcher Rentenversicherungsträger für eine Person zuständig ist, fehlt in den Daten.</p> <p>Seit der Einbeziehung der geringfügig Beschäftigten in das Meldeverfahren zum 01.04.1999 werden auch Entgelte unterhalb der Geringfügigkeitsgrenze erfasst, die Beitragsbemessungsgrenze als obere Kapfungsgrenze gilt jedoch weiterhin. In einigen Fällen übersteigen die gemeldeten Entgelte die Beitragsbemessungsgrenze allerdings trotzdem. Dies dürfte in der Regel auf Jahressonderzahlungen zurückzuführen sein, die der Arbeitgeber bei den Jahres-, Unterbrechungs- oder Abmeldungen zum normalen Entgelt hinzurechnen kann. Dabei ist es dann unerheblich, ob durch diese Addition die für den Meldezeitraum maßgebliche Beitragsbemessungsgrenze in der Rentenversiche-</p>

	<p>Angabe überschritten wird. Es könnte sich allerdings auch um fehlerhafte Angaben bei Beschäftigungszeit oder Entgelthöhe handeln, was wegen der Versicherungsrelevanz dieser Informationen jedoch eher selten der Fall sein dürfte. Aus datenschutzrechtlichen Gründen werden alle Entgelte, die oberhalb der Beitragsbemessungsgrenze (in vollen Euro) liegen, nicht mit ihrem tatsächlichen Wert ausgewiesen sondern durch den im Jahr der Meldung gültigen Euro Betrag der Beitragsbemessungsgrenze ersetzt. Die Geringfügigkeits- und Beitragsbemessungsgrenzen unterscheiden sich von Jahr zu Jahr und in West- und Ostdeutschland (maßgeblich ist jeweils der Standort des Betriebs). Eine Übersicht über diese Grenzen findet sich unter http://fdz.iab.de.</p> <p>Eine Angabe von 0 Euro ist neben den o.g. abgerundeten Entgelten unterhalb von 0,5 Euro auch auf "Unterbrechungsmeldungen" zurückzuführen. In diesen Zeiten besteht das Beschäftigungsverhältnis zwar rechtlich fort, es wird aber kein Entgelt gezahlt. Dies ist der Fall bei Krankheit nach Ende der Lohnfortzahlung, Mutterschaftszeiten und Sabbaticals.</p> <p>2) LeH</p> <p>Bei LeH-Datensätzen enthält das Merkmal den täglichen Leistungssatz, jeweils umgerechnet und gerundet auf volle Euro. Leistungssätze unter 0,5 Euro wurden dabei auf 0 Euro abgerundet.</p> <p>Es ist zu berücksichtigen, dass der tägliche Leistungssatz bei Datensätzen mit einem Original-Beginndatum, das vor dem 1.1.1998 liegt, für Werktage und bei Datensätzen mit einem Original-Beginndatum ab 1.1.1998 für Kalendertage gilt.</p> <p>Die Angabe von 0 Euro kann entweder durch die Abrundung sehr geringer Leistungssätze unterhalb von 0,5 Euro verursacht werden oder seit dem 1.1.2005 auch im Zusammenhang mit Sperrzeiten oder Leistungsunterbrechungen stehen. Ist für einen Satz mit Leistungssatz gleich 0 ein Abmeldegrund/ Beendigungsgrund/ Abgabegrund angegeben, so handelt es sich um eine Unterbrechungsmeldung. Bei Sätzen, die eine Sperrzeit widerspiegeln, ist der Anspruch derselbe wie vor Beginn der Sperrzeit.</p>
--	--

5.4.3 Beruf – ausgeübte Tätigkeit (beruf_gr)

Variablenlabel	Beruf - ausgeübte Tätigkeit, vergrößert
Variablenname	beruf_gr
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine

Detailbeschreibung	<p>1) BeH Die Berufsbezeichnung der im Meldezeitraum vom Beschäftigten ausgeübten Tätigkeit ist Bestandteil der vom Arbeitgeber übermittelten „Angaben zur Tätigkeit“.</p> <p>Der Arbeitgeber verschlüsselt die Tätigkeit des Beschäftigten hierbei gemäß der „Klassifizierung der Berufe. Systematisches und alphabetisches Verzeichnis der Berufsbenennungen“ (Hrsg. von der Bundesanstalt für Arbeit, Nürnberg 1988), in dem ca. 25.000 Berufsbezeichnungen zu finden sind. Die Berufsordnung besteht aus einer 3-stelligen Kennziffer und umfasst rund 330 Ausprägungen. Treffen mehrere Berufsbezeichnungen mit verschiedenen Schlüsselzahlen für einen Beschäftigten zu, ist der Arbeitgeber gehalten, diejenige Berufsbezeichnung zu wählen, welche die überwiegend ausgeübte Tätigkeit kennzeichnet (vgl. BA 2005, S. V).</p> <p>Im Rahmen der Anonymisierung des SIAB-Regionalfiles werden die aus der BeH stammenden Berufsangaben zu 120 Berufsaggregaten zusammengefasst.</p> <p>2) (X)ASU Bei der Anonymisierung wurde diese Variable für (X)ASU-Datensätze auf missing (.z) gesetzt.</p>
--------------------	--

5.4.4 Stellung im Beruf und Arbeitszeit (stib)

Variablenlabel	Stellung im Beruf und Arbeitszeit
Variablenname	stib
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Die berufliche Stellung des Beschäftigten im Meldezeitraum wird vom Arbeitgeber im Rahmen der „Angaben zur Tätigkeit“ übermittelt.</p> <p>Das Merkmal „Berufsstellung“ unterscheidet zunächst zwischen Vollzeit- und Teilzeitbeschäftigten; maßgeblich ist dabei das Verhältnis zwischen der arbeitsvertraglich vereinbarten und der betriebsüblichen Arbeitszeit. Nur für Vollzeitbeschäftigte liefert die Variable tatsächlich Angaben zur beruflichen Stellung, während für Teilzeitbeschäftigte nur erfasst wird, ob ihre Stundenzahl eine bestimmte Grenze übersteigt oder nicht. Diese Grenze lag bis 1978 bei 20 Wochenstunden, zwischen 1979 und 1987 bei 15 Wochenstunden und seit 1988 bei 18 Wochenstunden.</p> <p>Die Unterscheidung zwischen vollzeitbeschäftigten Arbeitern (1,2) und Angestellten richtet sich ausschließlich nach der Art der Rentenversicherungspflicht (BfA für Angestellte und LVA für Arbeiter). Neben den Auszubildenden, Praktikanten und Volontären zählen auch Anlernlinge, Schüler an Schulen des Gesundheitswesens und Teilnehmer an geförderten Maßnahmen zur beruflichen Fortbildung, Umschulung und betrieblichen Einarbeitung zur Kategorie Beschäftigte in beruflicher Ausbildung (0).</p> <p>Falls mehrere Schlüsselzahlen in Betracht kommen, ist der Arbeitgeber dazu angehalten, die Einstufung nach der überwiegend ausgeüb-</p>

	ten Tätigkeit vorzunehmen. Lässt sich diese nicht eindeutig feststellen, ist die Schlüsselzahl der höherwertigen Stellung im Beruf einzutragen (vgl. Bundesagentur für Arbeit (2005): S. VI).
--	---

5.4.5 Erwerbsstatus (erwstat_gr)

Variablenlabel	Erwerbsstatus vergrößert
Variablenname	erwstat_gr
Kategorie	Informationen zu Beschäftigung, Leistungsbezug und Arbeitsuche
Herkunft	BeH, LeH, (X)ASU
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal nimmt für jede Datenquelle unterschiedliche Werte mit unterschiedlichen Bedeutungen an.</p> <p>1) BeH Für BeH-Sätze ist die Personengruppe der Beschäftigtenmeldung enthalten, die mit dem neuen Meldeverfahren zum 1.1.1999 eingeführt wurde. Sie weist beitrags- oder leistungsrechtliche Besonderheiten des Beschäftigungsverhältnisses aus, die mindestens einen Sozialversicherungsträger betreffen, unterschieden nach meldenden Institutionen. Sozialversicherungspflichtige Arbeitnehmer, deren Beschäftigungsverhältnis keine Besonderheiten aufweist, erhalten den Schlüssel 101. Treffen mehrere Schlüssel zu, ist der kleinste anzugeben. Geringfügig Beschäftigte lassen sich in der SIAB als die Personengruppen 109 und 209 abgrenzen. Das Merkmal Personengruppe ist erst seit der Einführung des neuen Meldeverfahrens zum 1.1.1999 Bestandteil der Meldungen zur Sozialversicherung. Die Personengruppe kann in Meldungen, die sich auf die Jahre vor 1999 beziehen, aber erst 1999 oder später eingegangen sind, enthalten sein. Für Meldungen, die vor 1999 eingegangen sind, wird versucht, anhand bestimmter Regeln mit Hilfe der Merkmale „Ausbildung“, „Stellung im Beruf und Arbeitszeit“ und „Beruf“ sowie weiterer Informationen eine Zuordnung zu den Personengruppen durchzuführen. In vielen Fällen ist das jedoch nicht eindeutig möglich. Die Meldepflicht für geringfügig Beschäftigte besteht seit 01.04.1999. Vorher sind keine Daten über geringfügig Beschäftigte enthalten.</p> <p>2) LeH Für LeH-Sätze ist die gruppierte Leistungsart angegeben, d.h. es ist angegeben, ob Arbeitslosengeld, Arbeitslosenhilfe, Unterhaltsgeld bezogen werden oder ob Beiträge zur privaten Pflegeversicherung von der BA übernommen werden. Die Ausprägungen dieses Merkmals wurden durch Zusammenfassung der Ausprägungen des Merkmals Leistungsart generiert.</p> <p>3) (X)ASU Bei (X)ASU ist der Status der Arbeitsuche enthalten. Eine Unterscheidung nach der Herkunft aus den Originalquellen ASU und XASU ist</p>

	<p>durch deren Zusammenfassung nicht mehr möglich. Aus der XASU wurden nur die Statusangaben „nicht arbeitslos, aber arbeitsuchend“ bzw. „arbeitsuchend und gleichzeitig arbeitslos“ in die SIAB integriert. Diese Angaben zum Status der Arbeitsuche wurden mit dem entsprechenden Status aus der ASU zu einer gemeinsamen Kategorie zusammengefasst.</p> <p>Als Ratsuchende werden Empfänger von Arbeitslosengeld (bzw. Arbeitslosengeld II), die Leistungen unter den erleichterten Voraussetzungen des §428 SGB III (bzw. des §65 Abs. 4 SGB II) beziehen, sowie Nichtleistungsempfänger ohne Vermittlungsbereitschaft i. S. des § 252 Abs. 8 SGB VI, wenn das 58. Lebensjahr vollendet wurde, geführt.</p>
--	---

5.4.6 Beschäftigungsdauer vor dem Alter von 17 (tage_jung)

Variablenlabel	Tage in Beschäftigung vor der Altersuntergrenze (Jahr des 17. Geburtstags)
Variablenname	tage_jung
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal gibt an, wie viele Tage eine Person vor dem Jahr ihres 17. Geburtstags beschäftigt war. Gezählt wird dabei nur die Hauptbeschäftigung (Stata Syntax: <code>quelle_gr==1 & level2==0</code>).

5.4.7 Beschäftigungsdauer nach dem Alter von 62 (tage_alt)

Variablenlabel	Tage in Beschäftigung nach der Altersobergrenze (Jahr des 62. Geburtstags)
Variablenname	tage_jung
Kategorie	Informationen zu Beschäftigung
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	Dieses Merkmal gibt an, wie viele Tage eine Person nach dem Jahr ihres 62. Geburtstags beschäftigt war. Gezählt wird dabei nur die Hauptbeschäftigung (Stata Syntax: <code>quelle_gr==1 & level2==0</code>).

5.5 Betriebsmerkmale

5.5.1 Wirtschaftszweig 73 – Wirtschaftsgruppe 73 (w73_gen_gr)

Variablenlabel	Wirtschaftszweig 73, vergrößert, ab 2003 generiert
Variablenname	w73_gen_gr
Kategorie	Betriebsmerkmale
Herkunft	BeH

Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	<p>Dieses Merkmal basiert auf der Wirtschaftszweigklassifikation WS73. WS73 steht für das "Verzeichnis der Wirtschaftszweige für die Statistik der Bundesanstalt für Arbeit, Ausgabe 1973". Mittels eines dreistelligen Zahlencodes werden 269 Wirtschaftsklassen unterschieden, wobei die erste Stelle dieses Codes die Wirtschaftsabteilung, insgesamt 10, definiert und die beiden ersten Stellen zusammen die jeweilige Gruppe, insgesamt 95, festlegen.</p> <p>Jedem Betrieb darf nur eine Kennziffer zugeordnet werden. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, muss die wirtschaftliche Haupttätigkeit ermittelt werden.</p> <p>Angaben zur WS73 liegen nur von 1975 bis 2002 vor. Die Variable enthält für diesen Zeitraum die Angaben aus den Meldungen basierend auf der Wirtschaftsklasse W73. Ab 2003 werden sämtliche Werte, basierend auf einer Heuristik, ergänzt. Dazu werden zunächst gültige Werte in die Zukunft fortgeschrieben. Ist eine Fortschreibung nicht möglich, so werden weitere fehlende Werte anhand von Umschlüsselungstabellen ersetzt. Zur Methodik und für Qualitätshinweise sei auf Eberle et al. (2011) verwiesen.</p> <p>Im Rahmen der Anonymisierung des SIAB-Regionalfiles werden die Wirtschaftszweigangaben zu 16 Aggregaten zusammengefasst.</p> <p>In einzelnen Regionen wurden Wirtschaftszweige, die in mindestens einem der geprüften Jahre unter der Mindestfallzahl für die faktische Anonymität lagen, anonymisiert und auf missing (.z) umgesetzt (siehe Anhang).</p>

5.5.2 Wirtschaftszweig 03 – Wirtschaftsgruppe 03 (w03_gr)

Variablenlabel	Wirtschaftszweig 03, vergrößert
Variablenname	w03_gr
Kategorie	Betriebsmerkmale
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	keine
Detailbeschreibung	<p>Dieses Merkmal weist den Wirtschaftszweig gemäß der Klassifikation WZ03 als 3-Steller aus. WZ03 steht für die „Klassifikation der Wirtschaftszweige Ausgabe 2003“ des Statistischen Bundesamtes (Hrsg.). Der Wirtschaftszweig nach der 03er Klassifikation ist erst seit 2003 gefüllt.</p> <p>Jedem Betrieb darf nur eine Kennziffer zugeordnet werden. Ist ein Betrieb in unterschiedlichen wirtschaftlichen Bereichen tätig, muss die wirtschaftliche Haupttätigkeit ermittelt werden.</p> <p>Im Rahmen der Anonymisierung des SIAB-Regionalfiles werden die aus der BeH stammenden Berufsangaben zu 18 Berufsaggregaten zusammengefasst.</p> <p>In einzelnen Regionen wurden Wirtschaftszweige, die in mindestens einem der geprüften Jahre unter der Mindestfallzahl für die faktische Anonymität lagen, anonymisiert und auf missing (.z) umgesetzt (siehe Anhang).</p>

5.6 Ortsangaben

5.6.1 Arbeitsort Region (ao_region)

Variablenlabel	Arbeitsort Region
Variablenname	ao_region
Kategorie	Ortsangaben
Herkunft	BeH
Datentyp	Numerisch
Hierarchie	Kreisregion Bundesland
Detailbeschreibung	<p>Kreisfreie Städte und Landkreise wurden zu 333 Kreisregionen (vgl. Abschnitt 7.5) aggregiert, sodass jede dieser Regionen (zum 31.12.2008) eine Mindesteinwohnerzahl von 100.000 Einwohnern aufweist.</p> <p>Die Ausgangsdaten enthalten den Kreis (kreisfreie Stadt oder Landkreis), in dem der Betrieb des Beschäftigten angesiedelt ist. Der 5-stellige Kreisschlüssel enthält in den ersten beiden Stellen den Schlüssel des Bundeslands, in der ersten bis dritten Stelle den Regierungsbezirk und in der ersten bis fünften Stelle den Kreis. Bei Ländern ohne Regierungsbezirk steht an der dritten Stelle eine Null.</p> <p>Um konsistente regionale Zuordnungen über den gesamten Beobachtungszeitraum sicherzustellen, wurden die zugrunde liegenden Kreisangaben auf den Gebietsstand vom 31.12.2008 umgeschlüsselt, d.h. in allen Kalenderjahren richtet sich die Zuordnung eines Betriebsortes zu einem Kreis nach den Grenzen, welche die Kreise zum 31.12.2008 hatten. Da sich die Grenzen der Kreise im Zeitverlauf geändert haben, würden ohne Gebietsstandsaktualisierung Fälle auftreten, bei denen die Kreiskennziffer des Betriebsortes wechselt, ohne dass dieser Betrieb seinen Standort verlagert hat.</p>

5.6.2 Pendlerstatus (pendler)

Variablenlabel	Pendlerstatus
Variablenname	pendler
Kategorie	Ortsangaben
Herkunft	BeH, LeH, (X)ASU, (X)LHG
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Dieses Merkmal gibt an, ob der aktuelle oder letzte Arbeitsort einer Person dem aktuellen Wohnort entspricht.</p> <p>Als Referenz für alle Quellen wird der aktuelle bzw. der Arbeitsort der Person aus der letzten zeitlich vorher liegenden BEH-Meldung (Hauptbeschäftigung) zugrunde gelegt und mit dem aktuellen Wohnort verglichen.</p> <p>Das Merkmal weist folgende Ausprägungen aus:</p> <p>0: Gemeinde des Wohnortes entspricht der Gemeinde des (letzten gefüllten) Arbeitsortes (kein Pendler)</p>

	<p>1: Gemeinde des Wohnortes entspricht nicht der Gemeinde des (letzten gefüllten) Arbeitsortes, beide Gemeinden gehören aber der gleichen Kreisregion an (intra-regionaler Pendler).</p> <p>2: Gemeinde des Wohnortes und Gemeinde des (letzten gefüllten) Arbeitsortes liegen in unterschiedlichen Kreisregionen (inter-regionaler Pendler)</p> <p>Das Merkmal ist gefüllt für: BEH: 1999-2008 LEH: 2003-2008 (X)ASU: 2000-2004 und 2007-2008 (X)LHG: 2007-2008</p> <p>Ist für eine Episode einer Person in den Originaldaten der Wohnort einer Quelle nicht gefüllt, so wird das Merkmal Pendlerstatus auf .z gesetzt. Ebenso enthält das Merkmal .z, wenn zwar der Wohnort der Person in einer Quelle gefüllt ist, jedoch im Erwerbsverlauf vorher noch kein BEH-Satz mit gefülltem Arbeitsort als Referenzort existiert.</p>
--	---

5.7 Generierte technische Merkmale

5.7.1 Quelle des Satzes (quelle_gr)

Variablenlabel	Quelle des Satzes, vergrößert
Variablenname	quelle_gr
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, (X)ASU, (X)LHG
Datentyp	Numerisch
Hierarchie	Keine
Detailbeschreibung	<p>Das Merkmal gibt die Datenquelle an. Dabei wurden die Originalquellen ASU und XASU sowie LHG und XLHG jeweils zu einer Quelle zusammengefasst. Eine Unterscheidung nach der Trägerart ist aus Datenschutzgründen somit nicht mehr möglich.</p> <p>Durch Korrekturen unterscheiden sich die ausgewiesenen Zeiträume der Quellen im Vergleich zur SIAB 7508. Informationen zur Aufbereitung entnehmen Sie bitte Kapitel 3.</p>

5.7.2 Satzzähler Konto (spell)

Variablenlabel	Satzzähler Konto
Variablenname	spell
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, (X)LHG, (X) ASU
Datentyp	Numerisch
Hierarchie	Keine
Anonymisierung	Keine

Detailbeschreibung	Der Satzzähler Konto nummeriert aufsteigend die Sätze einer Person. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze. Mit dem Merkmal Satzzähler Konto kann die ursprüngliche Sortierung auf einfache Art und Weise wieder hergestellt werden. Die Sortierung erfolgt grundsätzlich zunächst nach dem Beginndatum der Episode und dann nach der Quelle.
--------------------	--

5.7.3 Satzzähler pro Episode (level2)

Variablenlabel	Satzzähler pro Episode
Variablenname	Level2
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, (X)ASU, (X)LHG
Datentyp	Numerisch
Hierarchie	Keine
Anonymisierung	Keine
Detailbeschreibung	<p>Zähler, der zeitlich parallel liegende Sätze zählt. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze.</p> <p>Dieser Levelzähler zeigt an, ob sich innerhalb eines Kontos Sätze zeitlich überschneiden. Dies ist beispielsweise bei mehreren gleichzeitig bestehenden Beschäftigungsverhältnissen oder beim Zusammentreffen von Leistungsbezug und geringfügiger Beschäftigung der Fall. level2 nummeriert die Sätze einer solchen Episode, beginnend bei 0. Innerhalb der Episode ist zunächst nach Quellen sortiert (BeH, LeH, (X)ASU, (X)LHG). Innerhalb der Quelle BeH stehen nicht-geringfügige Beobachtungen vor geringfügigen, danach ist absteigend nach Tagesentgelt sortiert. Innerhalb der Quelle LeH ist nach Leistungsart sortiert, innerhalb der anderen Quellen ist die Sortierung zufällig.</p>

5.7.4 Satzzähler pro Episode und Quelle (level1)

Variablenlabel	Satzzähler pro Episode und Quelle
Variablenname	level1
Kategorie	Generierte technische Merkmale
Herkunft	BeH, LeH, (X)ASU, (X)LHG
Datentyp	Numerisch
Hierarchie	Keine
Anonymisierung	Keine
Detailbeschreibung	<p>Mit Hilfe dieser Level-Variablen werden (zeitgleiche) Sätze pro Episode und Quelle gezählt. Das Merkmal wird beim Episodensplitting erzeugt und bezieht sich auf die gesplitteten Sätze.</p> <p>Dieser Levelzähler zeigt an, ob sich innerhalb eines Kontos Sätze</p>

	<p>zeitlich überschneiden. Level1 differenziert anders als Level2, der zeitlich parallele Sätze zählt, zusätzlich auch noch nach der Quelle dieser Sätze. Die Level1-Variable somit nimmt jeweils für den ersten Datensatz, der je Episode und Quelle vorliegt den Wert 0 an. Weitere parallel verlaufende Datensätze für dieselbe Episode und Quelle werden dann hochgezählt.</p> <p>Innerhalb der Quelle BeH stehen nicht-geringfügige Beobachtungen vor geringfügigen, danach ist absteigend nach Tagesentgelt sortiert. Innerhalb der Quelle LeH ist nach Leistungsart sortiert, innerhalb der anderen Quellen ist die Sortierung zufällig. (Für jede Quelle gilt: $level2 - level1 \equiv const$; für Quelle BeH ist $const = 0$.)</p>
--	---

6 Literaturverzeichnis

Bundesagentur für Arbeit (Hg.) (2005): Schlüsselverzeichnis für die Angaben zur Tätigkeit in den Meldungen zur Sozialversicherung. Ausgabe Januar 2005, Nürnberg.

Bundesagentur für Arbeit (Hg.) (1988): Klassifizierung der Berufe – Systematisches und Alphabetisches Verzeichnis der Berufsbenennungen, Nürnberg.

Bender, Stefan; Hilzendegen, Jürgen; Rohwer, Götz; Rudolph, Helmut (1996): Die IAB-Beschäftigtenstichprobe 1975-1990. Beiträge zur Arbeitsmarkt- und Berufsforschung 197, Nürnberg.

Christians, Helga; Wirth, Heike (2009): Mikrozensus-Regionalfile: Hochrechnung, Stichprobenfehler und Anonymisierung. Düsseldorf, Mannheim.

URL:http://www.forschungsdatenzentrum.de/bestand/mikrozensus_regionalfile/fdz_1_mz-regionalfile_dokumentation.pdf (Abrufdatum: 01.07.2011).

Cramer, Ulrich (1985): Probleme der Genauigkeit der Beschäftigtenstatistik. In: Allgemeines Statistisches Archiv 69: S. 56-68.

Dorner, Matthias; Heining, Jörg; Jacobebbinghaus, Peter; Seth, Stefan (2010): Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975-2008. (FDZ Datenreport, 01/2010 (de)), Nürnberg, 63 S.

Dorner, Matthias; Heining, Jörg; Jacobebbinghaus, Peter; Seth, Stefan (2010): The Sample of Integrated Labour Market Biographies. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 130, H. 4, S. 599-608.

Drews, Nils (2006): Qualitätsverbesserung der Bildungsvariable in der IAB-Beschäftigtenstichprobe 1975-2001. (FDZ Methodenreport, 05/2006 (de)), Nürnberg, 16 S.

Drews, Nils (2008): Das Regionalfile der IAB-Beschäftigtenstichprobe 1975-2004 - Handbuch-Version 1.0.2. (FDZ Datenreport, 02/2008 (de)), Nürnberg, 87 S.

Eberle, Johanna; Jacobebbinghaus, Peter; Ludsteck, Johannes; Witter, Julia (2011): Generation of time-consistent industry codes in the face of classification changes. Simple heuris-

tic based on the Establishment History Panel (BHP). (FDZ Methodenreport, 05/2011 (en)), Nürnberg, 21 S.

Fitzenberger, Bernd; Osikominu, Aderonke; Völter, Robert (2006): Imputation rules to improve the education variable in the IAB employment subsample. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 126, H. 3, S. 405-436.

Hethey-Maier, Tanja; Seth, Stefan (2010): Das Betriebs-Historik-Panel (BHP) 1975-2008 – Handbuch Version 1.0.2. (FDZ Datenreport, 04/2010 (de)), Nürnberg, 82 S.

Meinken, Holger; Koch, Iris (2004): BA-Beschäftigtenpanel 1998-2002. Codebuch, Nürnberg.

Müller, Walter; Blien, Uwe; Knoche, Peter; Wirth, Heike, et al. (1991): Die faktische Anonymität von Mikrodaten; Band 19 der Schriftreihe Forum der Bundesstatistik herausgegeben vom Statistischen Bundesamt; Stuttgart.

Oberschachtsiek, Dirk; Scioch, Patrycja; Seysen, Christian; Heining, Jörg (2009): Stichprobe der Integrierten Erwerbsbiografien IEBS - Handbuch für die IEBS in der Fassung 2008. (FDZ Datenreport, 03/2009 (de)), Nürnberg, 59 S.

Wermter, Winfried; Cramer, Ulrich (1988): Wie hoch war der Beschäftigtenanstieg seit 1983? – Ein Diskussionsbeitrag aus der Sicht der Beschäftigtenstatistik der Bundesanstalt für Arbeit. In: Mitteilungen aus der Arbeitsmarkt – und Berufsforschung 4/88, 468-482.

Zimmermann, Ralf; Kaimer, Steffen; Oberschachtsiek, Dirk (2007): Dokumentation des "Scientific Use Files der Integrierten Erwerbsbiographien" (IEBS-SUF V1) Version 1.0. (FDZ Datenreport, 01/2007 (de)), Nürnberg, 62 S.

7 Anhang

7.1 Unterschiede in der Anonymisierung zentraler Merkmale im Vergleich zur IABS-R04

	SIAB-Regionalfile 7508	IABS-Regionalfile 04
Angaben zum Wohnort	Indirekte Ausweisung durch Pendlerstatus (pendler).	Keine gültigen Wohnortangaben (region).
Geburtsjahre/ Altersangaben	Nur Meldungen mit Alter 17 bis 62 enthalten (gebjahr). Zusatzmerkmale weisen Beschäftigungszeiten vor 17 bzw. nach 62 Jahren aus (tage_alt, tage_jung).	Meldungen mit Alter 16 bis 62 ausgewiesen, Konten mit Beobachtungen unter bzw. über der Altersgrenze ohne gültige Altersinformationen (gebjahr).
Tagesentgelt/tägliche Leistungssätze	Ausweisung von gerundeten Entgelten aus Beschäftigung und Leistungssätzen, Beschäftigungsentgelte oberhalb der Beitragsbemessungsgrenze durch den Wert der Beitragsbemessungsgrenze ersetzt (tentgelt_gr)	Abgeschnittene ganzzahlige Tagesentgelte. Keine gültigen Angaben zu tägl. Leistungssätzen, Werte oberhalb der Beitragsbemessungsgrenze auf 999 bzw. unterhalb der Geringfügigkeitsgrenze auf 998 (tentgelt)

Tabelle A1: Unterschiede in der Anonymisierung zentraler Merkmale im Vergleich zur IABS-R04

7.2 Zusätzliche Bereinigungsverfahren im IABS-Regionalfile (IABS-R04) im Vergleich zum SIAB-Regionalfile (SIAB-R 7508)

In Abschnitt 1.3 wurden bereits die Unterschiede in der Aufbereitung von SIAB, IABS und IEBS kurz skizziert. Im Folgenden sollen die im Vergleich zur Erstellung der SIAB zusätzlich durchgeführten Aufbereitungsschritte bei der Generierung der IABS im Detail aufgezeigt werden.

Bereinigungen in der IABS 1975-2004
<p>Sätze gelöscht, falls Abgabegrund == 35 (Arbeitskampf)</p> <p>Sätze gelöscht, falls Abgabegrund 33 und Entgelt == 0</p> <p>Sätze gelöscht, falls Abgabegrund fehlt (nur LeH)</p>

Korrektur bei mehrfachem Leistungsbezug

Wenn eine Meldung vollständig innerhalb einer anderen, längeren Meldung liegt, wird die längere Meldung behalten und die kürzere gelöscht.

Wenn zwei Meldungen gleich lang sind, wird die Meldung mit der höchsten Priorität behalten:

Bei gleicher Priorität wird die Meldung mit dem höheren Leistungssatz behalten.

Bei Teilüberschneidung wird die zeitlich frühere Meldung gekürzt.

Korrektur der Überschneidungen von einem Tag

Es werden alle Fälle korrigiert, die sich genau um einen Tag überschneiden. Bei der Korrektur wird folgendermaßen vorgegangen:

- Erste Meldung hat Entgelt größer Null:
Erste Meldung wird komplett behalten, alle weiteren Meldungen der Episode werden um einen Tag verkürzt.
- Erste Meldung hat Entgelt gleich Null:
Sind in der Episode nur BeH-Meldungen, haben auf Grund der Sortierung aller Entgelte gleich Null und der erste Satz wird behalten. Darauf folgende Meldungen werden um einen Tag verkürzt. Sind in der Episode noch Meldungen zu geringfügiger Beschäftigung vorhanden oder LeH-Meldungen, dann wird die erste dieser Meldungen komplett behalten. Alle anderen werden korrigiert.

Die Korrektur wird nicht durchgeführt bei Meldungen mit Originaldauer von einem Tag.

Bereinigung bei Überschneidungen von BeH- und LeH - Meldungen

Hierbei gilt, dass in der Regel die BeH-Meldung die LeH- Meldung schlägt.

Ausnahmen sind BeH-Meldungen mit einem Entgelt gleich Null und BeH-Meldungen, die komplett im Zeitraum einer LeH-Meldung liegen. In diesen Fällen wird nur die BeH-Meldung gelöscht bzw. korrigiert.

Episoden, in denen nur Überschneidungen von BeH-Meldungen zu geringfügiger Beschäftigung und LeH-Meldungen vorliegen, werden nicht korrigiert.

Wird jedoch ein Satz gelöscht, werden alle weiteren Fälle der gleichen Episode und der gleichen Quelle, also auch geringfügige Beschäftigungsmeldungen gelöscht

Streikbereinigung

Wegen Aussperrung im Streik von Betrieben der Automobilindustrie in Baden-Württemberg und Hessen im Juni/Juli 1984 ergeben sich Beschäftigungsunterbrechungen. Daher wird für diese Personen eine Streikbereinigung durchgeführt. Es werden alle Sätze gelöscht, die in der Lücke liegen.

Ergänzungsverfahren

In der IABS werden Sätze ergänzt, wenn folgende Bedingungen erfüllt sind:

- Lücke zwischen zwei Datensätzen ist größer als sieben Tage
- Satz vor der Lücke, bzw. letzter Datensatz im Konto einer Person beschreibt:
 - Sozialversicherungspflichtige Beschäftigung (nicht geringfügig) , weist einen
 - Gültigen Eintrag im Merkmal blh_ao_kreis auf und
 - der Abgabegrund weist auf Beschäftigung hin (keine Abmeldung)
- Satz nach der Lücke ist eine BeH – Meldung.
- Jahresmeldungen am Ende eines Kontos werden nicht ergänzt.

Die Länge des ergänzten Zeitraumes hängt davon ab, ob der Zeitraum ins alte oder ins neue Meldeverfahren fällt. Ist die Lücke kleiner als der maximal zu ergänzende Zeitraum, wird der Zeitraum komplett ergänzt.

Für alle ergänzten Meldungen werden das Entgelt, das Tagesentgelt und der Abgabegrund auf 0 gesetzt, das Episodenanfangs- und Enddatum werden entsprechend der zu ergänzenden Zeiträume gesetzt. Die restlichen Merkmale werden von der Meldung vor der Lücke übernommen.

7.3 Übersicht Datumsanpassungen und Löschungen

	Beginn- und Enddatum Satz	Datumsangabe nach Zensierung	BeH	LeH	(X)ASU	(X)LHG
			1975-2008	1975-2008	2000-2004; 2007-2008	2007-2008
Löschungen	Beginn- und Enddatum vor 1.1.1975	-	X	X		
	Beginn- und Enddatum zwischen 1.1.2005 und 31.12.2006	-			X	
	Beginn- und Enddatum vor 1.1.2000	-			X	
	Beginn- und Enddatum vor 1.1.2007	-				X
	Beginn- und Enddatum nach 31.12.2008	-		X	X	X
Linkszenzierungen	Beginndatum vor 1.1.1975, Enddatum nach 1.1.1975	Beginndatum 1.1.1975		X		
	Beginndatum vor 1.1.2000, Enddatum nach 1.1.2000	Beginndatum 1.1.2000			X	
	Beginndatum zwischen 1.1.2005 und 31.12.2006, Enddatum nach 1.1.2007	Beginndatum 1.1.2007			X	X
Rechtszenzierung	Beginndatum vor 31.12.2008, Enddatum nach 31.12.2008	Enddatum 31.12.2008		X	X	X
	Beginndatum vor 31.12.2004, Enddatum zwischen 1.1.2005 und 31.12.2006	Enddatum 31.12.2004			X	

Links- u. Rechts- zensierung	Beginndatum vor 31.12.2004, Enddatum nach 31.12.2006	Enddatum Satz 1 31.12.2004; Beginndatum Satz 2 1.1.2007			X	
---------------------------------	---	--	--	--	---	--

Tabelle A2: Übersicht Datumsanpassungen und Löschungen bei der Aufbereitung des SIAB-Regionalfiles 7508

7.4 Vergleich SIAB 7508 und SIAB-Regionalfile 7508

Ausgangsdatei: Anzahl Personen SIAB 7508		1659024	100.0%	
1	Löschen von Sätzen in den Quellen BeH und LeH, die vor 1975 enden	1659024	100.0%	0
2	Löschen von Sätzen in den Quellen LeH,(X)ASU und (X)LHG, die nach 2008 beginnen	1652618	99,6 %	-0,4 %
3	Löschen von ASU-Sätzen, die vor 2000 enden	1557912	93,9 %	-5,7 %
4	Löschen von LHG, XASU, XLHG- Sätzen, die vor 2007 enden	1555976	93,7 %	-0,2 %
5	Löschen von ASU- Sätzen, die nach 2004 beginnen und vor 2007 enden	1552083	93,5 %	-0,2 %
6	Löschen von Sätzen in allen Quellen, die vor Alter 17 enden oder die nach Alter 62 beginnen	1515463	91,3 %	-2,2 %
Anzahl Personen SIAB-Regionalfile 7508		1515463	91,3 %	- 8,7%

Tabelle A3: Reduktion der Personenzahl im Zuge der Aufbereitung des SIAB-Regionalfiles

7.5 Umschlüsselungstabellen vergrößerter Merkmale im Vergleich zur schwach anonymisierten Version der SIAB 7508

Quelle des Satzes, vergrößert (quelle_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Merkmalsausprägung
1	1	BeH Beschäftigten-Historik
2	2	LEH Leistungsempfänger-Historik
7,18	3	(X)ASU kombinierte Arbeitsuchenden-Historik
16,19	4	(X)LHG kombinierte Leistungs-Historiken Grundsicherung

Tabelle A4: Merkmalsumschlüsselung „Quelle des Satzes, vergrößert“ (quelle_gr)

Erwerbsstatus, Leistungsart, Status der Arbeitsuche (erwstat_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Merkmalsausprägung
101, 140, 143	1	Sozialversicherungspflichtig Beschäftigte ohne bes. Merkmale
102, 141	2	Auszubildende
109, 209	3	Geringfügig Beschäftigte
103, 142	4	Beschäftigte in Altersteilzeit
105, 106	5	Praktikanten und Werkstudenten
118, 205	6	Unständig Beschäftigte
112, 119, 120, 201, 203	7	Sonstiger Beschäftigungsstatus
1	11	Arbeitslosengeld
2	12	Arbeitslosenhilfe
3	13	Unterhaltsgeld
5	14	Beitrag §166b AFG/§ 207a SGB III
31	21	Arbeitsuchend und gleichzeitig arbeitslos
32	22	Bis 6-wöchige Krankheit/Arbeitsunfähigkeit während der Arbeitslosigkeit
33	23	Nicht arbeitslos, aber arbeitsuchend
34	24	Ohne Status
35	25	Ratsuchend

Tabelle A5: Merkmalsumschlüsselung „Erwerbsstatus, Leistungsart, Status der Arbeitsuche, vergrößert“ (erwstat_gr)

Abgabegrund/ Beendigungsgrund/ SGB-II-Einstellungsgrund/ Abmeldegrund (grund_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Merkmalsausprägung
30, 49 vor 1999 auch: 31,32, 34, 35, 36, 40	1	Abmeldung wg. Ende Beschäftigung (auch Tod)
ab 1999: 31	2	Abmeldung wg. Krankenkassenwechsel
ab 1999: 32	3	Abmeldung wg. Beitragsgruppenwechsel
ab 1999: 34, 35	4	Abmeldung wg. Unterbrechung der Beschäftigung >1 Monat (auch Arbeitskampf)
ab 1999: 36	5	Abmeldung wg. Wechsel im Engeltabrechnungssystem
ab 1999: 40	6	Gleichzeitige Ab- u. Anmeldung wg. Ende der Beschäftigung
33, 54, 55, 56, 57, 59	7	Abmeldung wg. sonstiger Gründe
50	8	Jahresmeldung Beschäftigung
51, 52, 53	9	Unterbrechungsmeldung
2002	10	Arbeitsaufnahme
2001, 2021, 2039, 2041, 2046, 2048, 2051, 2052, 2053, 2056	11	Anspruch auf andere Leistung
2026, 2034, 2038	12	Sonstige Gründe
2008, 2015, 2036, 2040, 2042, 2055, 2057, 2058	13	Steht dem Arbeitsmarkt nicht zur Verfügung
2060	14	Ablauf der Maßnahme
2023, 2043, 2045, 2059	15	Sperr- und Säumniszeiten
2028	16	Anspruch erschöpft
2020	17	Bewilligungsabschnitt ausgelaufen
2003, 2004, 2006, 2007, 2010, 2011, 2012, 2014, 2016, 2017, 2019, 2022, 2024, 2027, 2029, 2031, 2032, 2033, 2035, 2037, 2044, 2047, 2049, 2050, 2054, 2061, 2062, 2063, 2064	18	Sonstige
5031, 5061	20	Vermittlung in ABM, SAM, BSI
5032, 5033, 5034, 5040	21	Vermittlung mit Eingliederungsvertrag (auch Be- hinderte), -zuschuss und Einstellungszuschuss bei Neugründungen
5035, 5036, 5079, 5088, 6012, 6027	22	Sonstige Maßnahmen der aktiven Arbeitsmarkt- politik
5039, 5071	23	FbW oder Deutschsprachlehrgang
5029, 5037, 5038, 5066, 6001	24	Beschäftigung (selbst gesucht)
5060, 5062, 5063, 5064, 5081, 5086, 5098, 6002, 6003, 6022, 6023, 6024	25	Vermittlung in Beschäftigung (>15 Std.) durch priv. Vermittler od. Arbeitsverwaltung, durch AG/Bewerber in Jobbörse gefunden
5030, 5048, 5082, 6005, 6021	26	Vermittlung in kurzfristige Beschäftigung, Be- schäftigung, gemeinnützige Tätigkeit, Erwerbstä- tigkeit unbekannt
5046, 5047, 5067, 6006	27	Selbständigkeit

5053, 5068, 6007, 6015, 6025	28	Wehr- und Zivildienst, Wehrübung, Soldat
5041, 6008	29	Erwerbs/Arbeitsunfähigkeit
5042, 5065, 6004	30	Rückruf/Wiedereinstellung bisheriger Arbeitgeber
5056	31	Verbleib im aktuellen Beschäftigungsverhältnis
5043, 5070	32	Betriebliche Ausbildung
5049, 5069, 6016	33	Schulische Ausbildung, Studium
5045, 5077	34	Umzug
5044, 5074, 6018, 6026	35	Fehlende Mitwirkung oder Verfügbarkeit
5052, 5076	36	Ausscheiden aus Erwerbsleben
5051, 5055, 6028	37	Beendigung der Hilfebedürftigkeit, vorübergehender Wegfall der Voraussetzungen
5073	38	Nichterneuerung der Meldung
5050, 5075, 5087, 6017, 6030, 6031	39	Sonderregelungen
5054, 5078, 5083, 6009, 6010, 6011, 6013, 6014, 6019, 6020, 6029	40	sonstige Gründe
5027	41	Sonstige weitere Leistungen
3003, 3001, 4004, 4017, 4018	50	Arbeitsaufnahme
3011, 3012, 3013, 4028, 4007, 4008, 4009	51	Wehr- u. Zivildienst, Wehrübung
4010, 3005	52	Übergangsgeld
4013, 3006	53	Ausbildung, Studium
4022, 4023, 3007	54	Umzug
4014, 3008	55	Eigene Abmeldung, Ortsabwesenheit ohne Leistungsgewährung
4029, 3009, 4006	56	Ende Gewährungszeitraum/Leistungsfortzahlung
4011, 4025, 3010, 3017, 4001	57	Altersrente, Grundsicherung im Alter
4002, 4012, 3016, 3004	58	(temp.) Wegfall der Erwerbsfähigkeit
4016, 4019, 4020, 4021, 4027, 3015	59	Wegfall der Bedürftigkeit
4003, 4024, 4026, 4015, 3014, 4005, 3002	60	Sonstige Gründe

Tabelle A6: Merkmalsumschlüsselung „Abgabegrund/Beendigungsgrund/SGB-II-Einstellungsgrund/ Abmeldegrund, vergrößert“ (grund_gr)

Berufe – ausgeübte Tätigkeit (beruf_gr)		
SIAB 7508	SIAB-R 7508 (SUF)	Merkmalsausprägung
011, 012, 013, 014, 015, 016, 017, 018, 019, 020, 021, 022, 023, 024, 025, 026, 027, 028, 029, 030, 031, 032, 041, 042, 043, 044	1	Landwirte bis Tierpfleger und verwandte Berufe
051, 052, 053, 054, 055, 056, 057, 058, 059, 060, 061, 062	2	Gärtner, Gartenarbeiter bis Waldarbeiter, Wald- nutzer
071, 072, 073, 074, 075, 076, 077, 078, 079, 080, 081, 082, 083, 084, 085, 086, 087, 088, 089, 090, 091, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112	3	Bergleute bis Formstein-, Betonhersteller
121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135	4	Keramiker bis Glasbearbeiter, Glasveredler
141	5	Chemiebetriebswerker
142, 143, 144	6	Chemielaborwerker bis Vulkaniseure
151	7	Kunststoffverarbeiter
161, 162, 163, 164	8	Papier-, Zellstoffhersteller bis sonstige Papierverarbeiter
171, 172, 173, 174	9	Schriftsetzer bis Flach-, Tiefdrucker
175, 176, 177	10	Spezialdrucker, Siebdrucker bis Druckerhelfer
181, 182, 183, 184	11	Holzaufbereiter bis Korb-, Flechtwarenmacher
191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203	12	Eisen-, Metallerzeuger, Schmelzer bis Halbzeugputzer und sonstige Formgießerberufe
211, 212, 213	13	Blechpresser, -zieher, -stanzer bis sonstige Metallverformer (spanlose Verformung)
221	14	Dreher
222, 223, 224	15	Fräser bis Bohrer
225, 226	16	Metallschleifer bis übrige spanende Berufe
231, 232, 233, 234, 235, 242, 243, 244	17	Metallpolierer bis Metallkleber und übrige Metall- verbinder
241	18	Schweißer, Brennschneider
251, 252, 263	19	Stahlschmiede bis Rohrnetzbauer, Rohrschlosser
261	20	Feinblechner
262	21	Rohrinstallateure
270, 271, 272	22	Schlosser o.n.a. bis Blech-, Kunststoffschlosser
273	23	Maschinenschlosser
274, 275	24	Betriebsschlosser, Reparaturschlosser bis Stahl- bauschlosser, Eisenschiffbauer
281	25	Kraftfahrzeuginstandsetzer
282, 283, 284	26	Landmaschineninstandsetzer bis Feinmechaniker
285, 286	27	Sonstige Mechaniker bis Uhrmacher
291, 301, 302	28	Werkzeugmacher bis Edelmetallschmiede

303, 304, 305, 306	29	Zahntechniker bis Puppenmacher, Modellbauer, Präperatoren
311	30	Elektroinstallateure, -monteure
312, 313, 315	31	Fernmeldemonteure, -handwerker bis Funk-, Tongerätemechaniker
314	32	Elektrogerätebauer
321	33	Elektrogeräte-, Elektroteilemontierer
322	34	Sonstige Montierer
323	35	Metallarbeiter o.n.a.
331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 371, 372, 373, 374, 375, 376, 377, 378	36	Spinner, Spinnvorbereiter bis Fellverarbeiter
351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362	37	Schneider bis Textilausrüster
391, 392	38	Backwarenhersteller bis Konditoren
401, 402, 403	39	Fleischer bis Fischverarbeiter
411, 412	40	Köche bis Fertiggerichte-, Obst-, Gemüsekonservierer, -zubereiter
421, 422, 423, 424, , 425, 426, 427, 428, 429, 430, 431, 432, 433	41	Weinküfer bis Zucker-, Süßwaren-, Speiseeishersteller
441, 442	42	Maurer bis Betonbauer
451, 453	43	Zimmerer bis Gerüstbauer
452	44	Dachdecker
461, 462	45	Pflasterer, Steinsetzer bis Strassenbauer
463, 464, 465, 466	46	Gleisbauer bis sonstige Tiefbauer
470, 471, 472	47	Bauhilfsarbeiter bis sonstige Bauhilfsarbeiter, Bauhelfer, a.n.g
481, 482	48	Stukkateure, Gipser, Verputzer bis Isolierer, Abdichter
483, 484, 485, 486	49	Fliesenleger bis Estrich-, Terazzoleger
491, 492, 502, 503, 504	50	Raumausstatter bis sonst. Holz-, Sportgerätebauer
501	51	Tischler
511	52	Maler, Lackierer (Ausbau)
512, 513, 514	53	Warenmaler, -lackierer bis Kerammaler, Glasma- ler
521	54	Warenprüfer, -sortierer, a.n.g.
522	55	Warenaufmacher, Versandfertigmacher
531	56	Hilfsarbeiter ohne nähere Tätigkeitsangabe
541, 542, 543, 544, 545, 546	57	Energiemaschinisten bis Baumaschinenführer
547, 548, 549	58	Maschinenwärter, Maschinistenhelfer bis Maschi- neneinrichter o.n.A.
601	59	Ingenieure des Maschinen- und Fahrzeugbaus
602	60	Elektroingenieure
603	61	Architekten, Bauingenieure
604, 605, 606, 607	62	Vermessungingenieure bis sonstige Ingenieure
611, 612	63	Chemiker, Chemieingenieure bis Physiker, Phy-

		sikingenieure, Mathematiker
621	64	Maschinenbautechniker
622, 623	65	Techniker des Elektrofaches bis Bautechniker
624, 625, 626, 627	66	Vermessungstechniker bis übrige Fertigungs- techniker
628	67	Sonstige Techniker
629	68	Industriemeister, Werkmeister
631, 632	69	Biologischtechnische Sonderfachkräfte bis physi- kalisch-, mathematisch-technische Sonderfach- kräfte
633, 634	70	Chemielaboranten bis Photolaboranten
635	71	Technische Zeichner
681	72	Gross- und Einzelhandelskaufleute, Einkäufer
682	73	Verkäufer
683, 684, 685, 686	74	Verlagskaufleute, Buchhändler bis Tankwarte
687, 688	75	Handelsvertreter, Reisende bis ambulante Händ- ler
691, 692	76	Bankfachleute bis Bausparkassenfachleute
693, 694	77	Krankenversicherungskaufleute (nicht Sozialver- sicherung) bis Lebens-, Sachversicherungskauf- leute
701	78	Speditionskaufleute
702, 703, 704, 705, 706	79	Fremdenverkehrsfachleute bis Geldeinnehmer, - auszahler, Kartenverkäufer, -kontrolleure
711, 712, 713, 715, 716	80	Schienenfahrzeugführer bis Strassenwarte
714	81	Kraftfahrzeugführer
721, 722, 723, 724, 725, 726	82	Nautiker bis Luftverkehrsberufe
731, 732, 733, 734	83	Posthalter bis Telefonisten
741	84	Lagerverwalter, Magaziner
742	85	Transportgeräteführer
743, 744	86	Stauer, Möbelpacker bis Lager-, Transportarbeiter
751	87	Unternehmer, Geschäftsführer, Geschäftsbe- reichsleiter
752, 753	88	Unternehmensberater, Organisatoren bis Wirt- schaftsprüfer, Steuerberater
761, 762, 763	89	Abgeordnete, Minister, Wahlbeamte bis Ver- bandsleiter, Funktionäre
771, 772	90	Kalkulatoren, Berechner bis Buchhalter
773	91	Kassierer
774	92	Datenverarbeitungsfachleute
781	93	Bürofachkräfte
782, 783	94	Stenographen, Stenotypisten, Maschinenschrei- ber bis Datentypisten
784	95	Bürohilfskräfte
791, 792	96	Werkschutzleute, Detektive bis Wächter, Aufse- her
793, 794	97	Pförtner, Hauswarte bis Haus-, Gewerbediener
801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812,	98	Soldaten, Grenzschutz-, Polizeibedienstete bis Rechtsvollstrecke

813, 814		
821, 822, 823	99	Publizisten bis Bibliothekare, Archivare, Museumsfachleute
831, 832, 833, 834	100	Musiker bis Dekorationen-, Schildermaler
835, 836, 837, 838	101	Künstlerische und zugeordnete Berufe der Bühnen-, Bild-, Tontechnik bis Artisten, Berufssportler, künstlerische Hilfsberufe
841, 842, 843, 844	102	Ärzte bis Apotheker
851, 852	103	Heilpraktiker bis Masseure, Krankengymnasten und verwandte Berufe
853	104	Krankenschwestern, -pfleger, Hebammen
854	105	Helfer in der Krankenpflege
855, 857	106	Diätassistenten, pharmazeutisch-technische Assistenten bis Medizinallaboranten
856	107	Sprechstundenhelfer
861, 863, 891, 892, 893	108	Sozialarbeiter, Sozialpfleger bis Seelsorge-, Kulturhelfer
862	109	Heimleiter, Sozialpädagogen
864	110	Kindergärtner, Kinderpfleger
871, 872, 873, 874	111	Hochschullehrer, Dozenten an höheren Fachschulen und Akademien bis Fachschul-, Berufsschul-, Werklehrer
875, 876, 877	112	Lehrer für musische Fächer bis sonstige Lehrer
881, 882, 883	113	Wirtschafts- und Sozialwissenschaftler, a.n.g., Statistiker bis Naturwissenschaftler a.n.g.
901, 902	114	Friseure bis sonstige Körperpfleger
911, 912	115	Gastwirte, Hoteliers, Gaststättenkaufleute bis Kellner, Stewards
913	116	Übrige Gästebetreuer
921, 922, 923, 924	117	Hauswirtschaftsverwalter bis mit Haushalts-scheckverfahren gemeldete Arbeitnehmer
931, 932	118	Wäscher, Plätter bis Textilreiniger, Färber und Chemischreiniger
933, 934	119	Raum-, Hausratreiniger bis Glas-, Gebäudereiniger
935, 936, 937	120	Strassenreiniger, Abfallbeseitiger bis Maschinen-, Behälterreiniger und verwandte Berufe

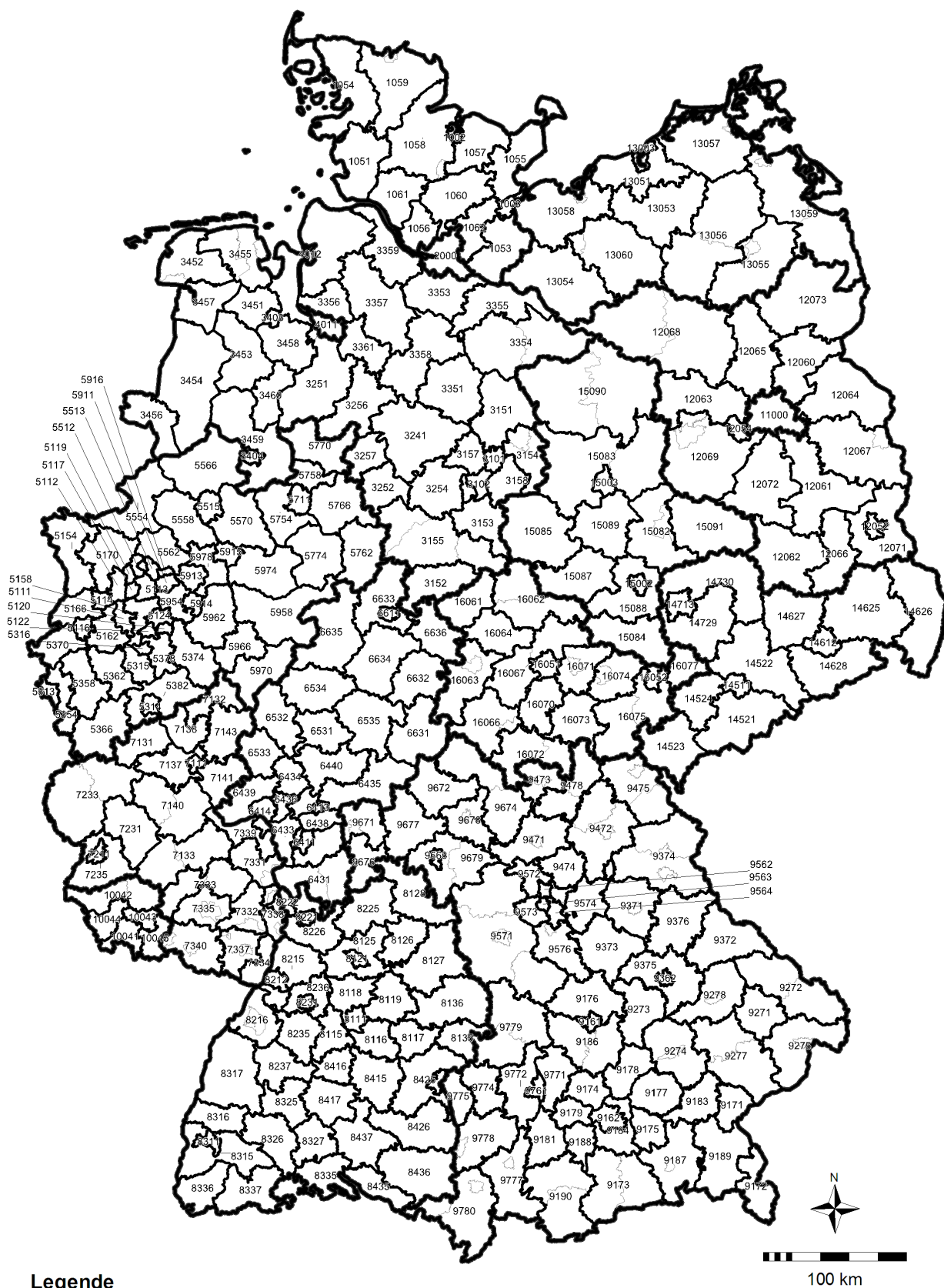
Tabelle A7: Merkmalsumschlüsselung „Beruf – ausgeübte Tätigkeit, vergrößert“ (beruf_gr)

Wirtschaftszweig 73 – Wirtschaftsgruppe 73, vergrößert (w73_gen_gr)			
SIAB 7508	SIAB-R 7508 (SUF)	Branchengruppe	Anonymisiert in den Kreisregionen
001-031, 040, 050-080	1	Landwirtschaft, Energie, Bergbau	3101, 3102, 3356, 6413, 7132, 8221, 8231, 9161, 9175, 9176, 9572, 9573, 9771, 9775
090-110, 130-146, 170-200, 220,221, 400,401, 430-433	2	Grundstoff-, Güterproduktion	1054, 12052
230-240, 260-300	3	Stahl- und Leichtmetallbau, Maschinenbau	
210, 211, 301-379	4	Stahlverformung, Fahrzeugbau, Gerätebau	3151, 9177
120, 150-162, 380-390, 410-421, 440-530	5	Konsumgütergewerbe	3151
540-581	6	Nahrungs- und Genussmittelgewerbe	3102, 5512, 6413, 12054
590-601	7	Baugewerbe (Haupt)	
250, 610-616	8	Baugewerbe (Ausbau)	
620, 621	9	Großhandel	9171, 9672
622-625, 850	10	Einzelhandel	
630-683	11	Verkehr und Nachrichtenübermittlung	
690, 691, 721, 774, 790-830, 851, 861, 862, 863, 865	12	Wirtschaftsbezogene Dienstleistungen	
700, 703, 720, 730, 731, 760-773	13	Haushaltsbezogene Dienstleistungen	
860, 864, 900	14	Heime, Krankenhäuser, Erziehung	
701, 702, 710, 711, 712, 740-758, 781-785	15	(Straßen)Reinigung, Verbände, Organisationen	
722, 780, 840-845, 870-890	16	Öffentliche Verwaltung, Sozialversicherung	5512

Tabelle A8: Merkmalsumschlüsselung „Wirtschaftszweig 73 – Wirtschaftsgruppe 73, vergrößert“ (w73_gen_gr)

Klassifikation der Wirtschaftszweige WS 2003, vergrößert“ (wz03_gr)			
SIAB 7508	SIAB-R 7508 (SUF)	Branchengruppe	Anonymisiert in folgenden Kreisregionen
001-014, 040, 041	1	Landwirtschaft, Energie, Bergbau	6413
015, 016	2	Nahrungs- und Genussmittelgewerbe	3102, 5512, 5916, 12054
017-022, 037	3	Bekleidung u.a., Holz, Papier, Verlagsgewerbe	3158, 13053
023-028, 036	4	Grundstoffproduktion, Recycling, sonst. Konsumgüter	
029-035	5	Maschinenbau, Fahrzeugbau, Feinmechanik	
045	6	Baugewerbe	
050	7	Kfz-Handel	
051	8	Großhandel, Handelsvermittlung	
052	9	Einzelhandel	
055	10	Gastgewerbe	
056-064	11	Verkehr und Nachrichtenübermittlung	
065-071	12	Kredit- und Versicherungsgewerbe, Grundstückswesen, Vermietungen	
072-074	13	Wirtschaftsbezogene Dienstleistungen, Forschung und Entwicklung, Datenverarbeitung	
075, 090, 096-098	14	Öffentliche Verwaltung, -Dienstleistungen	
080	15	Erziehung	
085	16	Gesundheits-, Veterinär- und Sozialwesen	
091, 092	17	Kirchliche Vereinigung, Kultur, Sport	
093-095	18	Sonstige Dienstleistungen, haushaltsbezogene Dienstleistungen	

Tabelle A9: Merkmalsumschlüsselung „Klassifikation der Wirtschaftszweige WS 2003, vergrößert“ (wz03_gr)



Legende

Bundesländer Deutschland
 Kreisregionen SIAB-Regionalfiler 7508
 krsfr. Städte/Landkreise SIAB 7508

Gebietsstand 31.12.2008; Kartographie: Matthias Dörner

Abbildung A1: Regionalisierung SIAB-Regionalfiler 7508 (ao_region)

Regionalisierung von kreisfreien Städten und Landkreisen, Gebietsstand 2008 (ao_region)		
SIAB 7508	SIAB-R 7508 (SUF)	Kreisregion
1002	1002	KS Kiel
1003	1003	KS Lübeck
1051	1051	Dithmarschen
1053	1053	Herzogtum Lauenburg
1054	1054	Nordfriesland
1055	1055	Ostholstein
1056	1056	Pinneberg
1057	1057	Plön
1058, 1004	1058	Rendsburg-Eckernförde, KS Neumünster
1059, 1001	1059	Schleswig-Flensburg, KS Flensburg
1060	1060	Segeberg
1061	1061	Steinburg
1062	1062	Stormarn
2000	2000	KS Hamburg
3101	3101	KS Braunschweig
3102	3102	KS Salzgitter
3151	3151	Gifhorn
3152	3152	Göttingen
3153	3153	Goslar
3154, 3103	3154	Helmstedt, KS Wolfsburg
3155, 3156, 3255	3155	Northeim, Holzminden, Osterode am Harz
3157	3157	Peine
3158	3158	Wolfenbüttel
3241	3241	Region Hannover
3251, 3401	3251	Diepholz, KS Delmenhorst
3252	3252	Hameln-Pyrmont
3254	3254	Hildesheim
3256	3256	Nienburg (Weser)
3257	3257	Schaumburg
3351	3351	Celle
3352, 3461	3352	Cuxhaven, Wesermarsch
3353	3353	Harburg
3354, 3360	3354	Lüchow-Dannenberg, Uelzen
3355	3355	Lüneburg
3356	3356	Osterholz
3357	3357	Rotenburg (Wümme)
3358	3358	Soltau-Fallingb.ostel
3359	3359	Stade
3361	3361	Verden
3403	3403	KS Oldenburg

3404	3404	KS Osnabrück
3451	3451	Ammerland
3452, 3402	3452	Aurich, KS Emden
3453	3453	Cloppenburg
3454	3454	Emsland
3455, 3405, 3462	3455	Friesland, Wittmund, KS Wilhelmshaven
3456	3456	Grafschaft Bentheim
3457	3457	Leer
3458	3458	Oldenburg
3459	3459	Osnabrück
3460	3460	Vechta
4011	4011	KS Bremen
4012	4012	KS Bremerhaven
5111	5111	KS Düsseldorf
5112	5112	KS Duisburg
5113	5113	KS Essen
5114	5114	KS Krefeld
5116	5116	KS Mönchengladbach
5117	5117	KS Mülheim an der Ruhr
5119	5119	Ks Oberhausen
5120	5120	Ks Remscheid
5122	5122	KS Solingen
5124	5124	KS Wuppertal
5154	5154	Kleve
5158	5158	Mettmann
5162	5162	Rhein-Kreis Neuss
5166	5166	Viersen
5170	5170	Wesel
5313	5313	KS Aachen
5314	5314	KS Bonn
5315	5315	KS Köln
5316	5316	KS Leverkusen
5354	5354	Aachen
5358	5358	Düren
5362	5362	Rhein-Erft-Kreis
5366	5366	Euskirchen
5370	5370	Heinsberg
5374	5374	Oberbergischer Kreis
5378	5378	Rheinisch-Bergischer Kreis
5382	5382	Rhein-Sieg-Kreis
5512	5512	KS Bottrop
5513	5513	KS Gelsenkirchen
5515	5515	KS Münster
5554	5554	Borken

5558	5558	Coesfeld
5562	5562	Recklinghausen
5566	5566	Steinfurt
5570	5570	Warendorf
5711	5711	KS Bielefeld
5754	5754	Gütersloh
5758	5758	Herford
5762	5762	Höxter
5766	5766	Lippe
5770	5770	Minden-Lübbecke
5774	5774	Paderborn
5911	5911	KS Bochum
5913	5913	KS Dortmund
5914	5914	KS Hagen
5915	5915	KS Hamm
5916	5916	KS Herne
5954	5954	Ennepe-Ruhr-Kreis
5958	5958	Hochsauerlandkreis
5962	5962	Märkischer Kreis
5966	5966	Olpe
5970	5970	Siegen-Wittgenstein
5974	5974	Soest
5978	5978	Unna
6411	6411	KS Darmstadt
6412	6412	KS Frankfurt am Main
6413	6413	KS Offenbach am Main
6414	6414	KS Wiesbaden
6431, 6437	6431	Bergstrasse, Odenwaldkreis
6432	6432	Darmstadt-Dieburg
6433	6433	Gross-Gerau
6434	6434	Hochtaunuskreis
6435	6435	Main-Kinzig-Kreis
6436	6436	Main-Taunus-Kreis
6438	6438	Offenbach
6439	6439	Rheingau-Taunus-Kreis
6440	6440	Wetteraukreis
6531	6531	Giessen
6532	6532	Lahn-Dill-Kreis
6533	6533	Limburg-Weilburg
6534	6534	Marburg-Biedenkopf
6535	6535	Vogelsbergkreis
6611	6611	KS Kassel
6631	6631	Fulda
6632	6632	Hersfeld-Rotenburg

6633	6633	Kassel
6634	6634	Schwalm-Eder-Kreis
6635	6635	Waldeck-Frankenberg
6636	6636	Werra-Meißner-Kreis
7111	7111	KS Koblenz
7131	7131	Ahrweiler
7132	7132	Altenkirchen (Westerwald)
7133, 7134	7133	Bad Kreuznach, Birkenfeld
7137	7137	Mayen-Koblenz
7138	7138	Neuwied
7140, 7135	7140	Rhein-Hunsrück-Kreis, Cochem-Zell
7141	7141	Rhein-Lahn-Kreis
7143	7143	Westerwaldkreis
7211	7211	KS Trier
7231	7231	Bernkastel-Wittlich
7233, 7232	7233	Daun, Bitburg-Prüm
7235	7235	Trier-Saarburg
7314	7314	KS Ludwigshafen am Rhein
7315	7315	Ks Mainz
7331, 7319	7331	Alzey-Worms, KS Worms
7332, 7316	7332	Bad Dürkheim, KS Neustadt an der Weinstraße
7333, 7336	7333	Donnersbergkreis, Kusel
7334	7334	Germersheim
7335, 7312	7335	Kaiserslautern, KS Kaiserslautern
7337, 7313	7337	Südliche Weinstrasse, KS Landau in der Pfalz
7338, 7311, 7318	7338	Rhein-Pfalz-Kreis, KS Speyer, KS Frankenthal (Pfalz)
7339	7339	Mainz-Bingen
7340, 7320, 7317	7340	Südwestpfalz, KS Zweibrücken, KS Pirmasens
8111	8111	KS Stuttgart
8115	8115	Böblingen
8116	8116	Esslingen
8117	8117	Göppingen
8118	8118	Ludwigsburg
8119	8119	Rems-Murr-Kreis
8121	8121	KS Heilbronn
8125	8125	Heilbronn
8126	8126	Hohenlohekreis
8127	8127	Schwäbisch Hall
8128	8128	Main-Tauber-Kreis
8135	8135	Heidenheim
8136	8136	Ostalbkreis
8212	8212	KS Karlsruhe
8215	8215	Karlsruhe

8216, 8211	8216	Rastatt, KS Baden-Baden
8221	8221	KS Heidelberg
8222	8222	KS Mannheim
8225	8225	Neckar-Odenwald-Kreis
8226	8226	Rhein-Neckar-Kreis
8231	8231	KS Pforzheim
8235	8235	Calw
8236	8236	Enzkreis
8237	8237	Freudenstadt
8311	8311	KS Freiburg im Breisgau
8315	8315	Breisgau-Hochschwarzwald
8316	8316	Emmendingen
8317	8317	Ortenaukreis
8325	8325	Rottweil
8326	8326	Schwarzwald-Baar-Kreis
8327	8327	Tuttlingen
8335	8335	Konstanz
8336	8336	Lörrach
8337	8337	Waldshut
8415	8415	Reutlingen
8416	8416	Tübingen
8417	8417	Zollernalbkreis
8421	8421	KS Ulm
8425	8425	Alb-Donau-Kreis
8426	8426	Biberach
8435	8435	Bodenseekreis
8436	8436	Ravensburg
8437	8437	Sigmaringen
9161	9161	KS Ingolstadt
9162	9162	KS München
9171	9171	Altötting
9172	9172	Berchtesgadener Land
9173, 9182	9173	Bad Tölz-Wolfratshausen, Miesbach
9174	9174	Dachau
9175	9175	Ebersberg
9176	9176	Eichstätt
9177	9177	Erding
9178	9178	Freising
9179	9179	Fürstenfeldbruck
9181	9181	Landsberg am Lech
9183	9183	Mühldorf a.Inn
9184	9184	München
9186, 9185	9186	Pfaffenhofen a.d.Ilm, Neuburg-Schrobenhausen
9187, 9163	9187	Rosenheim, KS Rosenheim

9188	9188	Starnberg
9189	9189	Traunstein
9190, 9180	9190	Weilheim-Schongau, Garmisch-Partenkirchen
9271	9271	Deggendorf
9272, 9276	9272	Freyung-Grafenau, Regen
9273	9273	Kelheim
9274, 9261	9274	Landshut, KS Landshut
9275, 9262	9275	Passau, KS Passau
9277, 9279	9277	Rottal-Inn, Dingolfing-Landau
9278, 9263	9278	Straubing-Bogen, KS Straubing
9362	9362	KS Regensburg
9371, 9361	9371	Amberg-Sulzbach, KS Amberg
9372	9372	Cham
9373	9373	Neumarkt i.d.Opf.
9374, 9363, 9377	9374	Neustadt a.d. Waldnaab, Tirschenreuth, KS Weiden i.d.Opf.
9375	9375	Regensburg
9376	9376	Schwandorf
9471, 9461	9471	Bamberg, KS Bamberg
9472, 9462, 9477	9472	Bayreuth, KS Bayreuth, Kulmbach
9473, 9463	9473	Coburg, KS Coburg
9474	9474	Forchheim
9475, 9464, 9479	9475	Hof, KS Hof, Wunsiedel i.Fichtelgebirge
9478, 9476	9478	Lichtenfels, Kronach
9562	9562	KS Erlangen
9563	9563	KS Fürth
9564	9564	KS Nürnberg
9571, 9561, 9577, 9575	9571	Ansbach, KS Ansbach, Neustadt a.d.Aisch-Bad Windsheim, Weißenburg-Gunzenhausen
9572	9572	Erlangen-Höchstadt
9573	9573	Fürth
9574	9574	Nürnberger Land
9576, 9565	9576	Roth, KS Schwabach
9663	9663	KS Würzburg
9671, 9661	9671	Aschaffenburg, KS Aschaffenburg
9672	9672	Bad Kissingen
9674, 9673	9674	Haßberge, Rhön-Grabfeld
9676	9676	Miltenberg
9677	9677	Main-Spessart
9678, 9662	9678	Schweinfurt, KS Schweinfurt
9679, 9675	9679	Würzburg, Kitzingen
9761	9761	KS Augsburg
9771	9771	Aichach-Friedberg
9772	9772	Augsburg
9774	9774	Günzburg

9775	9775	Neu-Ulm
9777, 9762	9777	Ostallgäu, KS Kaufbeuren
9778, 9764	9778	Unterallgäu, KS Memmingen
9779, 9773	9779	Donau-Ries, Dillingen a.d.Donau
9780, 9763, 9776	9780	Oberallgäu, KS Kempten (Allgäu), Lindau (Bodensee)
10041	10041	Stadtverband Saarbrücken
10042, 10046	10042	Merzig-Wadern, St. Wendel
10043	10043	Neunkirchen
10044	10044	Saarlouis
10045	10045	Saarpfalz-Kreis
11000, 11100, 11200	11000	KS Berlin (Ost und West)
12052	12052	KS Cottbus
12054	12054	KS Potsdam
12060	12060	Barnim
12061	12061	Dahme-Spreewald
12062	12062	Elbe-Elster
12063	12063	Havelland
12064	12064	Märkisch-Oderland
12065	12065	Oberhavel
12066	12066	Oberspreewald-Lausitz
12067, 12053	12067	Oder-Spree, KS Frankfurt (Oder)
12068, 12070	12068	Ostprignitz-Ruppin, Prignitz
12069, 12051	12069	Potsdam-Mittelmark, KS Brandenburg an der Havel
12071	12071	Spree-Neisse
12072	12072	Teltow-Fläming
12073	12073	Uckermark
13003	13003	KS Rostock
13051	13051	Bad Doberan
13053	13053	Güstrow
13054	13054	Ludwigslust
13055, 13002	13055	Mecklenburg-Strelitz, KS Neubrandenburg
13056, 13052	13056	Müritz, Demmin
13057, 13005, 13061	13057	Nordvorpommern, KS Stralsund, Rügen
13058, 13006	13058	Nordwestmecklenburg, KS Wismar
13059, 13001, 13062	13059	Ostvorpommern, KS Greifswald, Uecker-Randow
13060, 13004	13060	Parchim, KS Schwerin
14511	14511	KS Chemnitz
14521	14521	Erzgebirgskreis
14522	14522	Mittelsachsen
14523	14523	Vogtlandkreis
14524	14524	Zwickau
14612	14612	KS Dresden
14625	14625	Bautzen

14626	14626	Görlitz
14627	14627	Meissen
14628	14628	Sächs. Schweiz-Osterzgebirge
14713	14713	KS Leipzig
14729	14729	Leipzig
14730	14730	Nordsachsen
15002	15002	KS Halle/Saale
15003	15003	KS Magdeburg
15082, 15001	15082	Anhalt-Bitterfeld, Dessau-Rosslau
15083, 15086	15083	Börde, Jerichower Land
15084	15084	Burgenland
15085	15085	Harz
15087	15087	Mansfeld-Südharz
15088	15088	Saalekreis
15089	15089	Salzland
15090, 15081	15090	Altmarkkreis Salzwedel, Stendal
15091	15091	Wittenberg
16051	16051	KS Erfurt
16052	16052	KS Gera
16061	16061	Eichsfeld
16062, 16065	16062	Nordhausen, Kyffhäuserkreis
16063, 16056	16063	Wartburgkreis, KS Eisenach
16064	16064	Unstrut-Hainich-Kreis
16066, 16054	16066	Schmalkalden-Meiningen, KS Suhl
16067, 16068	16067	Gotha, Sömmerda
16070	16070	Ilm-Kreis
16071, 16055	16071	Weimarer Land, KS Weimar
16072, 16069	16072	Sonneberg, Hildburghausen
16073	16073	Saalfeld-Rudolstadt
16074, 16053	16074	Saale-Holzland-Kreis, KS Jena
16075, 16076	16075	Saale-Orla-Kreis, Greiz
16077	16077	Altenburger Land

Tabelle A10: Regionalisierung SIAB-Regionalfiler 7508 (ao_region)

7.6 Auszählungen

Auszählungen sowie Übersichten über die einzelnen Ausprägungen und Labels der Merkmale finden sich in gesonderten Dateien unter <http://fdz.iab.de>.

8 Abkürzungsverzeichnis

AA	Agentur für Arbeit / Arbeitsamt
ALG	Arbeitslosengeld
ARGE	Arbeitsgemeinschaft
ASU	Arbeitsuchenden-Historik
BA	Bundesagentur für Arbeit
BeH	Beschäftigten-Historik
BfA	Bundesversicherungsanstalt für Angestellte
BHP	Betriebs-Historik-Panel
BMAS	Bundesministerium für Arbeit und Soziales
DEÜV	Verordnung über die Erfassung und Übermittlung von Daten für die Träger der Sozialversicherung – Datenerfassungs- und -übermittlungsverordnung –
DEVO	Zweite VO über die Erfassung von Daten für die Träger der Sozialversicherung und für die BA – Datenerfassungs-Verordnung –
DÜVO	Zweite VO über die Datenübermittlung auf maschinell verwertbaren Datenträgern im Bereich der Sozialversicherung und der BA – Datenübermittlungs-Verordnung –
FDZ	Forschungsdatenzentrum der Bundesagentur für Arbeit am IAB
IAB	Institut für Arbeitsmarkt- und Berufsforschung
IABS	IAB-Beschäftigtenstichprobe
IEB	Integrierte Erwerbsbiografien
IEBS	Stichprobe der Integrierten Erwerbsbiografien
LeH	Leistungsempfänger-Historik
LHG	Leistungs-Historik Grundsicherung
LVA	Landesversicherungsanstalt
MTH	Maßnahmeteilnehmer-Historik
SGB	Sozialgesetzbuch
SIAB	Stichprobe der Integrierten Arbeitsmarktbiografien
SUF	Scientific Use File
VerBIS	Vermittlungs- und Beratungsinformationssystem
XASU	Arbeitsuchenden-Historik aus XSozial-BA-SGB II
XLHG	Leistungsempfänger-Historik Grundsicherung aus XSozial-BA-SGB II

Impressum

FDZ-Datenreport 7/2011

Herausgeber

Forschungsdatenzentrum (FDZ)
der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Str. 104
90478 Nürnberg

Redaktion

Stefan Bender, Iris Dieterich

Technische Herstellung

Iris Dieterich

Rechte

Nachdruck – auch auszugsweise – nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2011/DR_07-11.pdf

Internet

<http://fdz.iab.de/>

Rückfragen zum Inhalt an:

Matthias Dorner,
Forschungsdatenzentrum des Institut für
Arbeitsmarkt- und Berufsforschung (IAB),
Regensburger Str. 104,
90478 Nürnberg
Telefon: 0911 / 179-8895
E-Mail: matthias.dorner@iab.de